Castro et al. EURASIP Journal on Wireless Communications and Networking 2012, 2012:109
http://jwcn.eurasipjournals.com/content/2012/1/109

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Performance of MIMO systems in measured indoor channels with transmitter noise

Paula M Castro, José P González-Coma, José A García-Naya* and Luis Castedo

Abstract

This study analyzes the impact of transmitter noise on the performance of *multiple-input multiple-output* (MIMO) systems with linear and nonlinear receivers and precoders. We show that the performance of MIMO linear and decision-feedback receivers is not significantly influenced by the presence of transmitter noise, which does not hold true in the case of MIMO systems with precoding. Nevertheless, we also show that this degradation can be greatly alleviated when the transmitter noise is considered in the MIMO precoder design. A MIMO testbed developed at the University of A Coruña has been employed for experimentally evaluating how much the transmitter noise impacts the system performance. Both the transmitter noise and the receiver noise covariance matrices have been estimated from a set of 260 indoor MIMO channel realizations. The impact of transmitter noise has been assessed in this realistic scenario.

1 Introduction

Practical implementations of wireless transmitters suffer from a large number of impairments such as quantization noise, sampling offset, phase noise, I/Q imbalance ... which can be classified into systematic and non-systematic effects. These impairments are normally ignored when *multiple-input multiple-output* (MIMO) signaling methods are designed. However, noise generated at the transmitter can significantly affect predicted performance in practical scenarios. For instance, the performance of a linear *zero-forcing* (ZF) MIMO detector affected by the transmitter impairments of a MIMO *orthogonal frequency-division multiplexing* (OFDM) hardware demonstrator is tested in [1], where it is shown how the performance achieved by such noisy systems suffers from a loss greater than 4dB for a *bit error rate* (BER) of 10^{-2}. More specifically, the impact of residual transmitter *radio-frequency* (RF) impairments on both MIMO channel capacity and receiver performance is analyzed in [2]. It has been demonstrated that *maximum-likelihood* (ML) and Max-log *a posteriori probability* (APP) MIMO detection suffer from a substantial performance loss under the presence of weak transmitter noise, whereas linear ZF receivers are much less affected.

In this study, the impact of residual non-deterministic, non-systematic transmit impairments on MIMO signaling methods not considered in [1,2] is studied. First of all, we focus on MIMO systems with either linear receivers [3-5] or linear precoders [4-8]. We will also focus on MIMO systems with nonlinear *decision feedback* (DF) receivers [9,10] and nonlinear *Tomlinson-Harashima* (TH) precoders [5,11,12]. Both DF receivers and TH precoders are widely used because of their good trade-off between performance and complexity. The modulo operator at the receiver of a MIMO system with TH precoding has also motivated the proposal of a more general MIMO precoding technique where the data signal superimposed with a perturbation signal inputs a linear filter at the transmitter. This scheme is referred to as *vector precoding* (VP) in the literature [5,13]. The optimum perturbation signal is found with a closest point search in a lattice. Despite its larger complexity, VP outperforms TH precoding.

Although the impact of transmitter noise has been evaluated over spatially-white Rayleigh channels in [14,15] and some preliminary results obtained from testbed measurements in an indoor scenario have been presented in [15,16], only our study evaluates the performance of all the aforementioned schemes affected by transmitter noise over measured indoor channels by means of the testbed developed by the University of A Coruña [17]. Contrary to previous studies, not only the

* Correspondence: jagarcia@udc.es
Department of Electronics and Systems, University of A Coruña, Facultad de Informatica, Campus de Elviña s/n, A Coruña 15071, Spain

channel coefficients are measured using such a testbed, but also a estimate of all the noise covariance matrices existing in practical MIMO systems and of the *signal-to-transmitter noise ratio* (STxNR) are obtained and plugged into the robust filter expressions to guarantee a complete reproducibility of a real indoor environment and the evaluation in terms of BER performance of the proposed schemes under such realistic conditions.

In this study, we show that noise generated at the transmitter significantly affects the performance of MIMO systems with precoding. On the other hand, the performance of both linear and DF MIMO receivers is more robust against the presence of transmitter noise (cf. findings in the FP6-IST project MASCOT on multiuser MIMO communication systems [18]). At a first glance, it seems natural that MIMO systems with precoding be more sensitive to transmitter noise since channel equalization is carried out by processing signals prior transmission. However, it will be also shown that the performance of precoded MIMO systems can be substantially improved if the transmitter noise is considered inside the precoder design.

This study is organized as follows. Section 2 introduces the MIMO signal model which takes into account the aforementioned transmit RF impairments. The design of MIMO linear receivers and precoders considering the presence of transmitter noise is addressed in Section 3, whereas Section 4 focuses on MIMO nonlinear receivers (decision feedback) and precoders (Tomlinson-Harashima and vector precoding). Section 5 briefly introduces the MIMO testbed used to obtain measurements of the noise and of the channel parameters necessary to reproduce a realistic, indoor scenario. Next, the performance of the MIMO systems analyzed in the previous sections is evaluated for this indoor scenario. Some concluding remarks are presented in Section 6. Finally, Appendix 1 details all derivations corresponding to the expressions of all linear and nonlinear filters derived including transmitter noise in the optimizations.

Vectors and matrices are denoted by lower-case bold, and capital-bold letters, respectively. We use E[•], tr(•), $(\bullet)^{\mathrm{T}}$, $(\bullet)^{\mathrm{H}}$, and $\| \bullet \|_2$ for expectation, trace of a matrix, transposition, conjugate transposition, and Euclidean norm, respectively. The ith element of a vector v is denoted by v_i.

2 Signal model with transmitter noise

Let us consider a narrowband MIMO communication system with N_{t} transmit and N_{r} receive antennas. When considering only the receiver noise, this system can be represented by the following discrete-time model

$$\mathbf{y}[n] = \mathbf{H}\mathbf{x}[n] + \boldsymbol{\eta}_r[n] \in \mathbb{C}^{N_{\mathrm{r}}}, \tag{1}$$

where $\mathbf{x}[n] \in \mathbb{C}^{N_{\mathrm{t}}}$ represents the transmitted signals, $\boldsymbol{\eta}_r[n] \in \mathbb{C}^{N_{\mathrm{r}}}$ is the noise vector introduced by the receivers (which will be referred to as the *Rx-noise*), $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{r}} \times N_{\mathrm{t}}}$ is the MIMO channel matrix, and $\mathbf{y}[n] \in \mathbb{C}^{N_{\mathrm{r}}}$ is the received signal vector. Note that we assume a block-fading channel, where $\mathbf{H}$ remains constant during the transmission of a data frame. Rx-noise is complex-valued, Gaussian-distributed with zero mean and covariance matrix $C_{\boldsymbol{\eta}_r} = E\left[\boldsymbol{\eta}_{\mathrm{r}}[n]\boldsymbol{\eta}_{\mathrm{r}}^{H}[n]\right]$, i.e., $\boldsymbol{\eta}_r[n] \sim \mathcal{N}_{\mathbb{C}}\left(0, C_{\boldsymbol{\eta}_r}\right)$. Transmit energy is constrained to a value

$$E_{tx} = \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{x}}), \tag{2}$$

where $\boldsymbol{C}_{\boldsymbol{x}}$ is the covariance matrix corresponding to the input symbols $\mathbf{x}[n]$. Accordingly, we define the *signal-to-receiver-noise ratio* (SRxNR) as

$$\mathrm{SRxNR} = \frac{E_{tx} N_{\mathrm{r}}}{\mathrm{tr}\left(C_{\boldsymbol{\eta}_{\mathrm{r}}}\right)}, \tag{3}$$

where we have assumed that the channel is normalized to have a mean Frobenius norm equal to $N_{\mathrm{t}}\, N_{\mathrm{r}}$.

When the residual impairments at the transmitter are taken into account, a more accurate model for the transmitted signal is

$$\boldsymbol{x}_t[n] = \boldsymbol{x}[n] + \boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_{\mathrm{t}}}, \tag{4}$$

where $\boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_{\mathrm{t}}}$ will be denoted as the *Tx-noise*. The subscript $_t$ will be used hereafter to denote signals affected by Tx-noise. This noise encompasses different practical effects, for example phase-noise. As explained in [2], Tx-noise is adequately modeled as an additive Gaussian noise since it results from the sum of a large number of residual transmit impairments. Tx-noise is assumed to be zero-mean with covariance matrix $\boldsymbol{C}_{\boldsymbol{\eta}_t}$ and the STxNR is defined as

$$\mathrm{STxNR} = \frac{E_{tx}}{\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_t}\right)}, \tag{5}$$

where E_{tx} is fixed and given by Equation (2). As an example, practical implementations of the IEEE 802.11 (WiFi) standard achieve STxNR values ranging from 22 dB to 32 dB (see [2] and references therein). The Tx-noise is also assumed to be statistically independent from the Rx-noise. Note that we use the STxNR instead of the transmitter *error vector magnitude* (EVM), which is another measure of the signal modulation quality (see [19]) that additionally considers systematic errors.

Equations (1) and (4) are also adequate to represent a multi-antenna, multiuser wireless system where a base station equipped with N antennas communicates with K single-antenna users. When considering the uplink, $N_t = K$ and $N_r = N$, whereas $N_t = N$ and $N_r = K$ for the downlink. Although the channel model is the same, multiuser systems impose certain constraints on the signal processing that can be carried out to recover the transmitted information [20]. Indeed, for the uplink, the base station collects the signals from all the users, so that they can be separated using conventional MIMO receiver methods. These methods, however, cannot be used for the downlink since users do not normally cooperate among them and, consequently, each receiver does not know the signals from the other receivers. Therefore, in such a case signal separation can only be carried out by resorting to precoding or to any other form of transmit processing.

3 MIMO linear receivers and precoders

Figure 1 shows the block diagram of a general MIMO system with linear transmit and receive processing. Notice that transmitter noise has been also included. We assume the transmitter sends a set of N_r data streams encoded into the vector of transmitted symbols $\boldsymbol{u}[n] \in \mathbb{A}^{N_r}$ where $\mathbb{A}$ represents the utilized modulation alphabet. For linear transmit and receive processing, one transmit filter, $\boldsymbol{F} \in \mathbb{C}^{N_t \times N_r}$, so-called *precoding filter*, and one receive filter, $\boldsymbol{G} \in \mathbb{C}^{N_r \times N_r}$, are used to recover the transmitted data. Thus, after linear precoding, the transmitted signal is

$$\boldsymbol{x}_t[n] = \boldsymbol{F}\boldsymbol{u}[n] + \boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_t}, \tag{6}$$

and at the other end, the signal at the output of the receive filter $\boldsymbol{G}$ is given by

$$\hat{\boldsymbol{u}}_t[n] = \boldsymbol{G}\boldsymbol{H}\boldsymbol{x}_t[n] + \boldsymbol{G}\boldsymbol{\eta}_r[n] \in \mathbb{C}^{N_r}. \tag{7}$$

This signal is the input to a symbol detector (represented by $Q(\bullet)$ in Figure 1) that maps $\hat{\boldsymbol{u}}_t[n]$ onto the modulation alphabet $\mathbb{A}$ and produces the estimates of the transmitted symbols, $\tilde{\boldsymbol{u}}[n]$.

3.1 MIMO linear receiver design with transmitter noise

Let us consider a simplified MIMO linear transmission system where the transmit filter is a weighted identity matrix, i.e., $\boldsymbol{F} = p\mathbf{I}$. Thus, all the effort to compensate for the channel spatial *inter-symbol interference* (ISI) is carried out by the receiver filter $\boldsymbol{G}$ (see [3,5,21]). Elaborating on the signal model, the symbols at the output of the receiver filter can be written as

$$\hat{\boldsymbol{u}}_t[n] = \hat{\boldsymbol{u}}[n] + \boldsymbol{G}\boldsymbol{H}\boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_t}, \tag{8}$$

where $\hat{\boldsymbol{u}}[n] = p\boldsymbol{G}\boldsymbol{H}\boldsymbol{u}[n] + \boldsymbol{G}\boldsymbol{\eta}_r[n]$ would be the received signal if there were no Tx-noise.

Next, let us define the error vector $\varepsilon_t[n] = \boldsymbol{u}[n] - \hat{\boldsymbol{u}}_t[n]$. The optimum receive filter $\boldsymbol{G}$ and transmit weight p will be those that minimize the *mean square error* (MSE) $\mathrm{E}\left[\|\varepsilon_t[n]\|_2^2\right]$ which, under the presence of Tx-noise, can be expressed as follows [3]

$$\mathrm{E}\left[\|\varepsilon_t[n]\|_2^2\right] = \mathrm{E}\left[\|\varepsilon[n]\|_2^2\right] + \mathrm{tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right). \tag{9}$$

Notice that $\mathrm{E}\left[\|\varepsilon[n]\|_2^2\right]$ is the cost function to be minimized in conventional *minimum mean square error* (MMSE) linear receiver design where the transmitter noise is not accounted for.

Similarly to the study in [5], it can be demonstrated that by restricting $p \in \mathbb{R}^+$ and having in mind the transmit-energy constraint $E_{tx} = \mathrm{tr}(\boldsymbol{C}_x)$, the minimization of the cost function given by Equation (9) produces the MMSE solution for the linear receiver given by

$$\begin{aligned} \boldsymbol{G}_{\mathrm{MMSE}} &= \left(\frac{\boldsymbol{C}_{\boldsymbol{u}}^{-1}}{p_{\mathrm{MMSE}}} + p_{\mathrm{MMSE}}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}, \\ p_{\mathrm{MMSE}} &= \sqrt{\frac{E_{\mathrm{tx}}}{\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right)}}, \end{aligned} \tag{10}$$

where $\boldsymbol{C}_{\boldsymbol{u}}$ is the covariance matrix of the transmitted symbols $\boldsymbol{u}[n]$.

From the MMSE design of the MIMO linear receiver, it is straightforward to obtain the expressions for the ZF MIMO linear receiver [5]: it is the limiting case when $\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{H}\right)/E_{\mathrm{tx}} \to 0$, i.e.,

$$\begin{aligned} \boldsymbol{G}_{\mathrm{ZF}} &= \left(p_{\mathrm{ZF}}\boldsymbol{H}^{H}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{H}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}, \\ p_{\mathrm{ZF}} &= \sqrt{\frac{E_{\mathrm{tx}}}{\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{u}})}}. \end{aligned} \tag{11}$$

Note that a necessary condition for the existence of the filter $\boldsymbol{G}_{\mathrm{ZF}}$ is that $N_t \le N_r$ since it exists only if $\boldsymbol{H}^{H}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}$ is invertible.

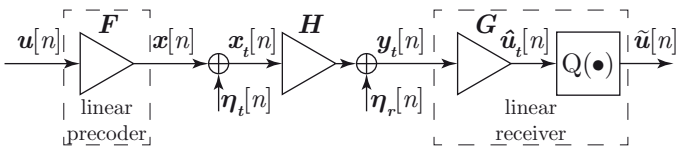

Figure 1 MIMO linear transmit processing and linear receive processing with Tx-noise.

3.2 MIMO linear precoder design with transmitter noise

Let us consider the MIMO transmission scheme dual to the previous one: the transmitted symbols $\boldsymbol{u}[n] \in \mathbb{C}^{N_r}$ are linearly precoded with the transmit filter $\boldsymbol{F} \in \mathbb{C}^{N_t \times N_r}$ and the receive filter is a weighted identity matrix $\boldsymbol{G} = g\mathbf{I}$, where $g \in \mathbb{R}^+$. Thus, the transmitted signal is

$$\boldsymbol{x}_t[n] = \boldsymbol{F}\boldsymbol{u}[n] + \boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_t}, \tag{12}$$

and the signal at the receiver input is given by

$$\boldsymbol{y}_t[n] = \boldsymbol{y}[n] + \boldsymbol{H}\boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_r}, \tag{13}$$

where $\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n] + \boldsymbol{\eta}_r[n]$ would be the received signal if there were no Tx-noise. After multiplication by the receive gain, g, we get the estimated symbols

$$\hat{\boldsymbol{u}}_t[n] = \hat{\boldsymbol{u}}[n] + g\boldsymbol{H}\boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_r}, \tag{14}$$

where $\hat{\boldsymbol{u}}[n] = g\boldsymbol{y}[n]$. We restrict the scalar value g to being common to all the receivers to simplify the optimization procedure.

The optimum transmit and receive filters are those that minimize the MSE between $\boldsymbol{u}[n]$ and $\hat{\boldsymbol{u}}_t[n]$ subject to the transmit-energy constraint $E_{tx} = \mathrm{tr}(\boldsymbol{C}_x)$ (see [3,5-7]). Again, let us define the error vector $\epsilon_t[n] = \boldsymbol{u}[n] - \hat{\boldsymbol{u}}_t[n]$ and express the MSE cost function as

$$\mathrm{E}\left[\left\|\varepsilon_t[n]\right\|_2^2\right] = \mathrm{E}\left[\left\|\varepsilon[n]\right\|_2^2\right] + \mathrm{tr}\left(|g|^2\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right), \tag{15}$$

where $\mathrm{E}\left[\left\|\varepsilon[n]\right\|_2^2\right]$ is the MSE when there is no Tx-noise [4,5,8].

Minimizing this cost function produces the MMSE solution for the MIMO linear precoding filter given by

$$\begin{aligned}\boldsymbol{F}_{\mathrm{MMSE}} &= g_{\mathrm{MMSE}}^{-1}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi_t\mathbf{I}\right)^{-1}\boldsymbol{H}^{\mathrm{H}},\\ g_{\mathrm{MMSE}} &= \sqrt{\frac{\mathrm{tr}\left(\left(\boldsymbol{H}^H\boldsymbol{H} + \xi_t\mathbf{I}\right)^{-2}\boldsymbol{H}^H\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}\right)}{E_{\mathrm{tx}}}},\end{aligned} \tag{16}$$

where $\xi_t = \xi + \mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)/E_{\mathrm{tx}}$, with ξ being the inverse of the SRxNR defined in Equation (5).

By applying the matrix inversion lemma to the MMSE solution of Equation (16) and considering afterwards the limiting case when $\xi_t \to 0$, the expressions for the ZF linear precoder can be readily obtained and expressed as follows

$$\begin{aligned}\boldsymbol{F}_{\mathrm{ZF}} &= g_{\mathrm{ZF}}^{-1}\boldsymbol{H}^H\left(\boldsymbol{H}\boldsymbol{H}^H\right)^{-1},\\ g_{\mathrm{ZF}} &= \sqrt{\frac{\mathrm{tr}\left(\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{C}_{\boldsymbol{u}}\right)}{E_{\mathrm{tx}}}}.\end{aligned} \tag{17}$$

4 MIMO nonlinear receivers and precoders

In this section, we focus on MIMO systems that use either nonlinear transmitters or nonlinear receivers to recover the transmitted data. It is well known that *maximum likelihood detection* (MLD) is the optimum detection scheme in the sense of minimizing the probability of a symbol being erroneously detected. The computational complexity of MLD grows exponentially with the number of transmit antennas and the modulation alphabet size and, for this reason, suboptimum nonlinear detection schemes such as the decision feedback one are preferred in a large number of practical situations.

However, decision-feedback receivers suffer from the major drawback of error propagation caused by feeding back erroneous decisions. One way to avoid this harmful effect is to perform a nonlinear filtering similar to that in DF but at the transmitter side. This idea leads to the concept of *Tomlinson-Harashima precoding* (THP). Finally, VP is another form of nonlinear MIMO transmit processing which will be considered in this section. Similarly to MLD, VP consists in a lattice search carried out at the transmitter instead of at the receiver side.

The impact of transmitter noise on the performance of MLD has already been analyzed in [2]. In the following sections, we will derive the expressions of the filters for the remaining nonlinear transceivers when the Tx-noise is accounted for.

4.1 MIMO decision feedback receiver design with transmitter noise

Figure 2 plots the block diagram of a MIMO system with a DF receiver. Information symbols will be represented by $\boldsymbol{u}[n] \in \mathbb{A}^{N_t}$, where $\mathbb{A}$ denotes the modulation alphabet, which are directly sent by the transmit antennas. It is assumed that $\boldsymbol{u}[n]$ is zero mean with covariance matrix denoted by $\boldsymbol{C}_u$.

It is apparent from Equations (1) and (4) that the input signal at the receiver can be written as

$$\boldsymbol{y}_t[n] = \boldsymbol{y}[n] + \boldsymbol{H}\boldsymbol{\eta}_t[n] \in \mathbb{C}^{N_r}, \tag{18}$$

where $\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{u}[n] + \boldsymbol{\eta}_r[n]$ is the received signal if there is no Tx-noise [5].

In DF reception, the signals at the channel output are passed through the feedforward filter $\boldsymbol{G} \in \mathbb{C}^{N_t \times N_r}$, which forces the ISI to be spatially causal and the error to be spatially white (i.e., minimum variance). By means of the feedback filter $\boldsymbol{I} - \boldsymbol{B} \in \mathbb{C}^{N_t \times N_t}$ and of the feedback loop shown in Figure 2, ISI can be recursively canceled without changing the statistical properties of the noise provided that the noise variance is sufficiently small so that the symbol detector (represented by Q(•) in Figure 2)

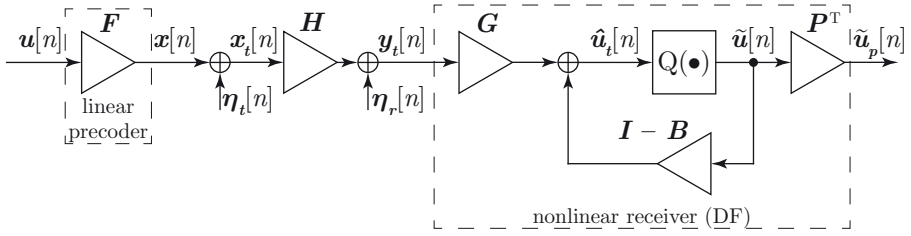

Figure 2 MIMO linear transmit processing and nonlinear receive processing (DF) with Tx-noise.

produces correctly detected symbols. By elaborating on the signal model according to Figure 2, the estimated signal $\hat{\boldsymbol{u}}_t[n]$ can be written as

$$\hat{\boldsymbol{u}}_t[n] = \boldsymbol{G}\boldsymbol{y}_t[n] + (\boldsymbol{I} - \boldsymbol{B})\tilde{\boldsymbol{u}}[n], \tag{19}$$

where $\boldsymbol{y}_t[n]$ is defined as in Equation (18) and $\tilde{\boldsymbol{u}}[n] \in \mathbb{A}^{N_t}$ denotes the detected symbols after the threshold quantizer.

The order in which symbols are detected has a significant influence on the performance of DF MIMO receivers. In the system model shown in Figure 2, the ordering is obtained with the multiplication of the detected symbols, $\tilde{\boldsymbol{u}}[n]$, by the permutation matrix $\boldsymbol{P}^{\mathrm{T}}$. This multiplication produces $\tilde{\boldsymbol{u}}_p[n]$, which constitutes the vector of detected symbols conveniently sorted. Having in mind that $\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}} = \mathbf{I}$, we have that $\tilde{\boldsymbol{u}}[n] = \boldsymbol{P}\tilde{\boldsymbol{u}}_p[n]$ and hence, $\hat{\boldsymbol{u}}_t[n]$ can be rewritten as

$$\hat{\boldsymbol{u}}_t[n] = \boldsymbol{G}\boldsymbol{y}_t[n] + (\boldsymbol{I} - \boldsymbol{B})\boldsymbol{P}\tilde{\boldsymbol{u}}_p[n].$$

The MMSE design of the DF MIMO receiver searches for the filtering and permutation matrices that minimize the variance of the error vector

$$\varepsilon_{t,p}[n] = \boldsymbol{P}\boldsymbol{u}[n] - \hat{\boldsymbol{u}}_t[n].$$

Assuming correct decisions (i.e., $\tilde{\boldsymbol{u}}_p[n] = \boldsymbol{u}[n]$) and according to Equation (18), this error vector can be rewritten as

$$\varepsilon_{t,p}[n] = \boldsymbol{B}\boldsymbol{P}\boldsymbol{u}[n] - \boldsymbol{G}\boldsymbol{y}_t[n] = \varepsilon_p[n] - \boldsymbol{G}\boldsymbol{H}\boldsymbol{\eta}_t[n],$$

where $\varepsilon_p[n] = \boldsymbol{B}\boldsymbol{P}\boldsymbol{u}[n] - \boldsymbol{G}\boldsymbol{y}[n]$ is the error vector when there is no Tx-noise. Since the Tx-noise is independent from the Rx-noise and the transmitted signals, the MSE cost function to be minimized can be written as

$$\mathrm{E}\left[\left\|\varepsilon_{t,p}[n]\right\|_2^2\right] = \mathrm{E}\left[\left\|\varepsilon_p[n]\right\|_2^2\right] + \mathrm{tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right), \tag{20}$$

where $\mathrm{E}\left[\left\|\varepsilon_p[n]\right\|_2^2\right]$ is the MSE with no Tx-noise. Notice that $\mathrm{E}\left[\left\|\varepsilon_p[n]\right\|_2^2\right]$ is the cost function that is minimized in the conventional MMSE design, whereas the additional term $\mathrm{tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right)$ is the MSE improvement caused by the inclusion of the Tx-noise.

An MMSE design of the MIMO link that accounts for the Tx-noise should minimize the MSE given by Equation (20). Similarly to the scenario without Tx-noise [22], minimization of Equation (20) is readily accomplished from the Cholesky factorization with symmetric permutation of

$$\boldsymbol{\Phi}_t = \left(\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1}\boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right)^{-1}.$$

This factorization is given by $\boldsymbol{P}\boldsymbol{\Phi}_t\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}$, where $\boldsymbol{L}$ is a unit lower triangular matrix and $\boldsymbol{D}$ is a diagonal matrix. After this decomposition, it can be demonstrated that the filters $\boldsymbol{G}$ and $\boldsymbol{B}$ for the MMSE DF nonlinear MIMO receiver solution are

$$\begin{aligned} \boldsymbol{G}_{\mathrm{MMSE}}^{\mathrm{DF}} &= \boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1}, \\ \boldsymbol{B}_{\mathrm{MMSE}}^{\mathrm{DF}} &= \boldsymbol{L}^{-1}. \end{aligned} \tag{21}$$

The minimum value of the MSE cost function is obtained plugging $\boldsymbol{G}_{\mathrm{MMSE}}^{\mathrm{DF}}$ and $\boldsymbol{B}_{\mathrm{MMSE}}^{\mathrm{DF}}$ into Equation (20). Hence, the MMSE value is

$$\mathrm{MMSE}_{t,\mathrm{DF}} = \mathrm{tr}(\boldsymbol{D}), \tag{22}$$

where $\boldsymbol{D}$ is the diagonal matrix obtained from the Cholesky factorization with symmetric permutation of $\boldsymbol{\Phi}_t$.

Notice that the MMSE expression given by Equation (22) depends on the permutation matrix $\boldsymbol{P}$. Brute force optimization of $\boldsymbol{P}$ can be carried out by computing the MMSE for all the $N_t!$ possible permutation matrices and choosing the one that provides the minimum value of Equation (22). Alternatively, more efficient ordering algorithms (such as the one described in [22]) can be used.

From the MMSE design of the DF receiver, it is straightforward to obtain the expressions for the ZF DF receiver: it is the limiting case when $\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)/E_{\mathrm{tx}} \to 0$. The final expressions for the ZF DF filters are exactly the same as before although $\boldsymbol{L}$ and $\boldsymbol{D}$ should be obtained from the Cholesky

decomposition of

$$\boldsymbol{\Phi}_t = \left(H^{\mathrm{H}}\left(HC_{\eta_t}H^{\mathrm{H}} + C_{\eta_r}\right)^{-1}H\right)^{-1}.$$

4.2 MIMO Tomlinson-Harashima precoder design with transmitter noise

Figure 3 shows the block diagram of a MIMO system employing THP. THP is a nonlinear precoding technique made up of a feedforward filter $\boldsymbol{F} \in \mathbb{C}^{N_t \times N_r}$, a feedbackward filter $\boldsymbol{I} - \boldsymbol{B} \in \mathbb{C}^{N_r \times N_r}$, and a modulo operator, represented in Figure 3 by M(•). The modulo operator is introduced to avoid the increase in transmit power due to the feedback loop [11]. Data symbols sent from the transmitter will be represented by $\boldsymbol{u}[n] \in \mathbb{A}^{N_r}$, where $\mathbb{A}$ denotes the modulation alphabet. The ordering considerably affects the performance of THP and, for this reason, transmit symbols are passed through a permutation filter $\boldsymbol{P}$. Minimization is carried out under the restriction of $\boldsymbol{B}$ being a spatially causal filter and E_{tx} being the transmitted energy, i.e., $\mathrm{E}\left[\|\boldsymbol{x}[n]\|_2^2\right] = E_{\mathrm{tx}}$, where $\boldsymbol{x}[n] = \boldsymbol{F}\boldsymbol{v}[n] \in \mathbb{C}^{N_t}$ is the transmitted signal, with $\boldsymbol{v}[n]$ representing the output of the modulo operator. At reception, we assume that all the receive antennas apply the same positive real value denoted by g. These assumptions are necessary in order to arrive at closed-form, unique solutions for the MMSE THP design.

In order to carry out the THP optimization, and taking into account the linear representation of THP [5,11], the desired signal is denoted by $\boldsymbol{d}[n]$ and it is expressed as

$$\boldsymbol{d}[n] = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v}[n]. \tag{23}$$

The received signal under the presence of Tx-noise is rewritten as

$$\hat{\boldsymbol{d}}_{\mathrm{t}}[n] = \hat{\boldsymbol{d}}[n] + g\boldsymbol{H}\boldsymbol{\eta}_{\mathrm{t}}[n] \in \mathbb{C}^{N_r}, \tag{24}$$

where $\hat{\boldsymbol{d}}[n] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{v}[n] + g\boldsymbol{\eta}_{\mathrm{r}}[n]$ is the received signal when there is no Tx-noise. At the receivers, the modulo operator is applied again to invert the effect of this operator at the transmitter and the resulting signal is passed through a symbol detector (represented by Q(•) in Figure 3) to produce the detected symbols $\tilde{\boldsymbol{u}}[n] \in \mathbb{A}^{N_r}$.

As explained in [5], the MMSE THP design searches for the filtering and permutation matrices that minimize the variance of the error vector

$$\varepsilon_{\mathrm{t}}[n] = \boldsymbol{P}^T\boldsymbol{B}\boldsymbol{v}[n] - \hat{\boldsymbol{d}}_{\mathrm{t}}[n] = \varepsilon[n] - g\boldsymbol{H}\boldsymbol{\eta}_t[n],$$

where $\boldsymbol{\varepsilon}[n] = \boldsymbol{P}^T\boldsymbol{B}\boldsymbol{v}[n] - g\boldsymbol{y}[n]$ is the error vector when there is no Tx-noise.

Since the Tx-noise is independent from the transmitted signal and the Rx-noise, the MSE can be decomposed as

$$\mathrm{E}\left[\|\varepsilon_t[n]\|_2^2\right] = \mathrm{E}\left[\|\varepsilon[n]\|_2^2\right] + |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right), \tag{25}$$

where $\mathrm{E}\left[\|\varepsilon[n]\|_2^2\right]$ is the MSE when there is no Tx-noise, which constitutes the cost function that is minimized in the conventional MMSE design of THP.

Following similar derivations as in [12], the minimization of the MSE cost function in Equation (25), subject to the mentioned constraints, can be carried out from the factorization of

$$\boldsymbol{\Phi}_t = \left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\boldsymbol{I}\right)^{-1},$$

where

$$\xi_{\mathrm{t}} = \xi + \mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)/E_{\mathrm{tx}}, \tag{26}$$

with $\xi = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r}\right)/E_{\mathrm{tx}}$. The symmetrically permuted Cholesky decomposition of this matrix is

$$\boldsymbol{P}\boldsymbol{\Phi}_{\mathrm{t}}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{L}, \tag{27}$$

where $\boldsymbol{L}$ and $\boldsymbol{D}$ are, respectively, unit lower triangular and diagonal matrices. Finally, the MMSE solution for the THP filters that account for the Tx-noise is given by

$$\begin{aligned}\boldsymbol{F}_{\mathrm{MMSE}}^{\mathrm{THP}} &= g_{\mathrm{MMSE}}^{\mathrm{THP},-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D},\\ \boldsymbol{B}_{\mathrm{MMSE}}^{\mathrm{THP}} &= \boldsymbol{L}^{-1}\end{aligned} \tag{28}$$

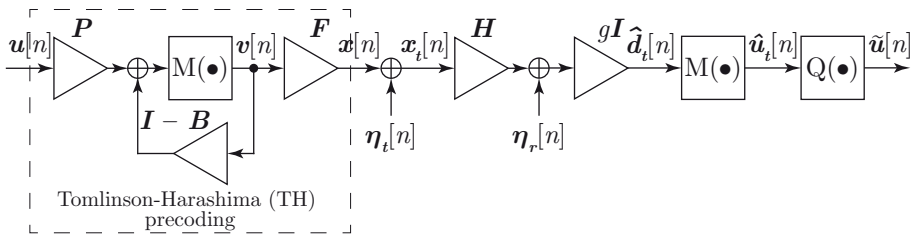

Figure 3 Nonlinear MIMO system with THP and Tx-noise.

The receive scalar weight $g_{\mathrm{MMSE}}^{\mathrm{THP}}$ is directly obtained from the transmit-energy constraint. Assuming that it is real and positive, it is obtained that

$$g_{\mathrm{MMSE}}^{\mathrm{THP}} = \sqrt{\frac{\mathrm{tr}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}^{2}\boldsymbol{C}_{v}\boldsymbol{L}\boldsymbol{P}\boldsymbol{H}\right)}{E_{\mathrm{tx}}}}, \tag{29}$$

where $\boldsymbol{C}_{v}$ is the covariance matrix of $\boldsymbol{v}[n]$, which is diagonal with entries depending on the modulation alphabet [11].

The minimum value for the MSE cost function given by Equation (25) can be obtained by substituting the expressions obtained for the optimum filters $\boldsymbol{F}_{\mathrm{MMSE}}^{\mathrm{THP}}$ and $\boldsymbol{B}_{\mathrm{MMSE}}^{\mathrm{THP}}$, and for the gain factor $g_{\mathrm{MMSE}}^{\mathrm{THP}}$. It is easy to show that the final MMSE under the presence of Tx-noise is

$$\mathrm{MMSE}_{t,\mathrm{THP}} = \xi_{\mathrm{t}}\mathrm{tr}\left(\boldsymbol{C}_{v}\boldsymbol{D}\right), \tag{30}$$

where ξ_{t} is given by Equation (26) and $\boldsymbol{D}$ is the diagonal matrix that results from the permuted Cholesky factorization of Equation (27).

As it is done in [12], instead of testing all the possible permutation matrices to find the one that minimizes the cost function of Equation (30), the ordering optimization can be included in the computation of the Cholesky decomposition of Equation (27).

Again, it is straightforward to obtain the expressions for the ZF THP design as the limiting case when $\xi_{\mathrm{t}} \to 0$. The expressions for the filters $\boldsymbol{F}_{\mathrm{ZF}}^{\mathrm{THP}}$ and $\boldsymbol{B}_{\mathrm{ZF}}^{\mathrm{THP}}$ are equal to those obtained for $\boldsymbol{F}_{\mathrm{MMSE}}^{\mathrm{THP}}$ and $\boldsymbol{B}_{\mathrm{MMSE}}^{\mathrm{THP}}$, respectively, although the matrices $\boldsymbol{P}$, $\boldsymbol{L}$, and $\boldsymbol{D}$ should be obtained from the symmetrically permuted Cholesky factorization of

$$\boldsymbol{\Phi}_{t} = \left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}.$$

4.3 MIMO vector precoder design with transmitter noise

Figure 4 shows the block diagram of a MIMO system with VP. The transmitter has the freedom to add an arbitrary perturbation signal $\boldsymbol{a}[n] \in \tau\mathbb{Z}^{N_{\mathrm{r}}} + \mathrm{j}\tau\mathbb{Z}^{N_{\mathrm{r}}}$ to the data signal prior to the linear transformation with the filter $\boldsymbol{F} \in \mathbb{C}^{N_{\mathrm{t}}\times N_{\mathrm{r}}}$. This perturbation will be later on removed by the modulo operator M(•) at the receiver. Here, τ denotes a constant that depends on the modulation alphabet.[a] This constant is associated with the nonlinear modulo operator M(•), defined as

$$\mathrm{M}(x) = x - \left(\left\lfloor \frac{\Re(x)}{\tau} + \frac{1}{2} \right\rfloor \tau + \mathrm{j}\left\lfloor \frac{\Im(x)}{\tau} + \frac{1}{2} \right\rfloor \tau\right) \in \mathbb{V}, \tag{31}$$

where $\lfloor \bullet \rfloor$ denotes the floor operator which gives the largest integer smaller than or equal to the argument. The corresponding fundamental Voronoi region is

$$\mathbb{V} = \left\{x \in \mathbb{C} | -\frac{\tau}{2} \le \Re(x) < \frac{\tau}{2}, -\frac{\tau}{2} \le \Im(x) < \frac{\tau}{2}\right\},$$

which means that the modulo operator constrains the real and imaginary part of x to the interval $[-\tau/2, \tau/2]$ by adding integer multiples of τ and j τ to the real and imaginary part, respectively.

As it can be seen from Figure 4, the data vector $\boldsymbol{u}[n] \in \mathbb{C}^{N_{\mathrm{r}}}$ is first superimposed with the perturbation vector $\boldsymbol{a}[n]$, and the resulting vector is then processed by the linear filter $\boldsymbol{F}$ to form the transmitted signal $\boldsymbol{x}[n] = \boldsymbol{F}\boldsymbol{d}[n] \in \mathbb{C}^{N_t}$, $n = 1,\ldots, N_{\mathrm{B}}$, where $\boldsymbol{d}[n]$ is the desired signal given by $\boldsymbol{u}[n] + \boldsymbol{a}[n]$ and n is the symbol index in a block size of N_{B} data symbols. The transmit-energy constraint is expressed as $\sum_{n=1}^{N_{\mathrm{B}}} \|\boldsymbol{x}[n]\|_2^2 / N_{\mathrm{B}} \le E_{\mathrm{tx}}$ since transmit-symbols statistics are unknown.

The weight g in Figure 4 is assumed to be constant throughout the block of N_{B} symbols. Again, note that a common weight for all the receivers is used. Thus, the weighted estimated signal is given by

$$\hat{\boldsymbol{d}}_{\mathrm{t}}[n] = \hat{\boldsymbol{d}}[n] + g\boldsymbol{H}\boldsymbol{\eta}_{\mathrm{t}}[n],$$

with $\hat{\boldsymbol{d}}[n] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{d}[n] + g\boldsymbol{\eta}_{\mathrm{r}}[n]$. The modulo operator at the receiver compensates the effect of adding the perturbation $\boldsymbol{a}[n]$ at the transmitter.

Since $\boldsymbol{a}[n]$ is discrete, their optimum values cannot be obtained after derivation. The optimization procedure is as follows. We start by fixing $\boldsymbol{a}[n]$, after which $\boldsymbol{x}[n]$ and g are optimized taking into account the transmit power constraint. For these optimum $\boldsymbol{x}[n]$ and g we choose the

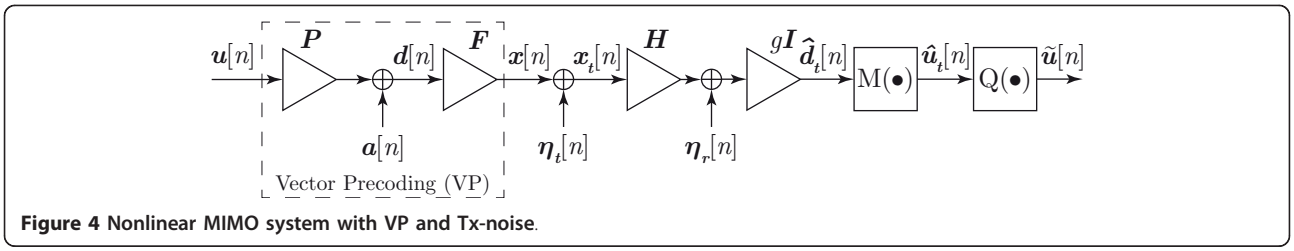

Figure 4 Nonlinear MIMO system with VP and Tx-noise.

best $\boldsymbol{a}[n]$ in order to minimize the following MSE criterion [13]

$$\mathrm{MSE}_{t,\mathrm{VP}} = \frac{1}{N_\mathrm{B}} \sum_{n=1}^{N_\mathrm{B}} \mathrm{E}\left[\left\| \boldsymbol{d}[n] - \hat{\boldsymbol{d}}_\mathrm{t}[n] \right\|_2^2 | \boldsymbol{u}[n] \right]. \tag{32}$$

With Tx-noise being present, the previous MSE cost function can be expressed as

$$\mathrm{MSE}_{t,\mathrm{VP}} = \mathrm{MSE}_\mathrm{VP} + \frac{1}{N_\mathrm{B}} \sum_{n=1}^{N_\mathrm{B}} \left\| \boldsymbol{g}\boldsymbol{H}\boldsymbol{\eta}_\mathrm{t}[n] \right\|_2^2, \tag{33}$$

where $\mathrm{MSE}_\mathrm{VP} = \sum_{n=1}^{N_\mathrm{B}} \mathrm{E}\left[\left\| \boldsymbol{d}[n] - \hat{\boldsymbol{d}}[n] \right\|_2^2 | \boldsymbol{u}[n] \right] / N_\mathrm{B}$ is the MSE cost function when the Tx-noise is away. Following an optimization procedure similar to that described in [13], we arrive at the MMSE VP solution given by

$$\begin{aligned} \boldsymbol{x}_\mathrm{MMSE}^\mathrm{VP}[n] &= \frac{1}{g_\mathrm{MMSE}^\mathrm{VP}} \left(\boldsymbol{H}^\mathrm{H}\boldsymbol{H} + \xi_\mathrm{t}\boldsymbol{I} \right)^{-1} \boldsymbol{H}^\mathrm{H}\boldsymbol{d}[n], \\ g_\mathrm{MMSE}^\mathrm{VP} &= \sqrt{\frac{\sum_{n=1}^{N_\mathrm{B}} \boldsymbol{d}^\mathrm{H}[n]\boldsymbol{H}\left(\boldsymbol{H}^\mathrm{H}\boldsymbol{H} + \xi_t\boldsymbol{I} \right)^{-2} \boldsymbol{H}^\mathrm{H}\boldsymbol{d}[n]}{E_\mathrm{tx} N_\mathrm{B}}}, \end{aligned} \tag{34}$$

where $g_\mathrm{MMSE}^\mathrm{VP}$ is directly obtained from the transmit-energy constraint.

By defining the matrix $\boldsymbol{\Phi}_\mathrm{t} = (\boldsymbol{H}\boldsymbol{H}^\mathrm{H} + \xi_\mathrm{t}\mathbf{I})^{-1}$ and applying the matrix inversion lemma to Equation (34), the MSE cost function given by Equation (33) is reduced to [13]

$$\mathrm{MMSE}_{t,\mathrm{VP}} = \frac{\xi_t}{N_\mathrm{B}} \sum_{n=1}^{N_\mathrm{B}} \boldsymbol{d}^\mathrm{H}[n]\boldsymbol{\Phi}_\mathrm{t}\boldsymbol{d}[n].$$

Since $\boldsymbol{\Phi}_\mathrm{t}$ is positive definite, we can use the Cholesky factorization to obtain a lower triangular matrix $\boldsymbol{L}$ and a diagonal matrix $\boldsymbol{D}$ with the following relationship,

$$\boldsymbol{\Phi}_\mathrm{t} = \left(\boldsymbol{H}\boldsymbol{H}^\mathrm{H} + \xi_\mathrm{t}\mathbf{I} \right)^{-1} = \boldsymbol{L}^\mathrm{H}\boldsymbol{D}\boldsymbol{L}.$$

Thus, the perturbation signal can be found by means of the following search [13]

$$\begin{aligned} \boldsymbol{a}_\mathrm{MMSE}^\mathrm{VP}[n] &= \underset{a[n] \in \tau\mathbb{Z}^{N_\mathrm{T}} + j\tau\mathbb{Z}^{N_\mathrm{T}}}{\operatorname{argmin}} \left(\boldsymbol{u}[n] + \boldsymbol{a}[n] \right)^\mathrm{H} \boldsymbol{\Phi}_\mathrm{t} \left(\boldsymbol{u}[n] + a[n] \right) \\ &= \underset{a[n] \in \tau\mathbb{Z}^{N_\mathrm{T}} + j\tau\mathbb{Z}^{N_\mathrm{T}}}{\operatorname{argmin}} \left\| \boldsymbol{D}^{1/2}\boldsymbol{L}\left(\boldsymbol{u}[n] + \boldsymbol{a}[n] \right) \right\|_2^2. \end{aligned} \tag{35}$$

This search can be solved using the Schnorr-Euchner sphere-decoding algorithm [23]. It is interesting to note that THP can be interpreted as a suboptimum approach to VP where $\boldsymbol{a}[n]$ is successively computed.

The ZF constraint $\mathrm{E}[\hat{\boldsymbol{d}}[n]|\boldsymbol{d}[n]] = \boldsymbol{g}\boldsymbol{H}\boldsymbol{F}\boldsymbol{d}[n]$, for $n = 1,\ldots, N_\mathrm{B}$, leads to similar expressions for $\boldsymbol{x}_\mathrm{ZF}^\mathrm{VP}[n]$ and $g_\mathrm{ZF}^\mathrm{VP}$ as those obtained for the MMSE VP design when $\xi_t \to 0$. Following similar steps as before, the cost function for ZF VP has exactly the same form as that of MMSE VP but considering $\boldsymbol{\Phi}_\mathrm{t} = (\boldsymbol{H}\boldsymbol{H}^\mathrm{H})^{-1}$. Finally, the optimum perturbation vectors are found by the following closest point search in a lattice

$$\boldsymbol{a}_\mathrm{ZF}^\mathrm{VP}[n] = \underset{a[n] \in \tau\mathbb{Z}^{N_\mathrm{T}} + j\tau\mathbb{Z}^{N_\mathrm{T}}}{\operatorname{argmin}} \left\| \boldsymbol{H}^\mathrm{H}\left(\boldsymbol{H}\boldsymbol{H}^\mathrm{H} \right)^{-1} \boldsymbol{d}[n] \right\|_2^2.$$

5 Experimental results

In this section, the results of several computer simulations are presented to illustrate the impact of Tx-noise on the performance of all the MIMO transmission systems already described in the previous sections. Figure 5 shows a picture of the transmitter of the wireless testbed developed at the University of A Coruña [17], which has been used to obtain realistic 4 × 4 MIMO channel realizations from an indoor scenario. We also measured the covariance matrices of both the Tx-noise and the Rx-noise corresponding to this specific testbed. The interaction with the testbed is solved by means of a custom-designed, distributed, multilayer software architecture [17].

5.1 Testbed details

As explained in [24,25], both transmit and receive testbed nodes are equipped with a Quad Dual-Band front-end from Lyrtech, Inc [26]. This RF front-end is equipped with up to eight antennas,[b] which are connected to four direct-conversion transceivers by means of an antenna switch. The front-end is based on Maxim [27] MAX2829 chip (also found in front-ends like Ettus [28] XCVR2450 or Sundance [29] SMT911). It supports both up and down conversion operations from either a 2.4 to 2.5 GHz band or a 4.9 to 5.875 GHz band. The front-end also incorporates a programmable variable attenuator to control the transmit power. The attenuation ranges from 0 to 31dB in 1 dB steps, while the

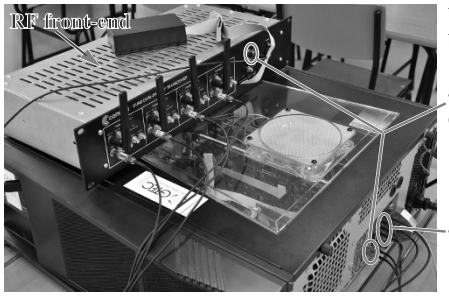

Figure 5 Picture of testbed transmitter developed at the University of A Coruña.

maximum transmit power declared by Lyrtech is 25dBm per transceiver.

The baseband hardware is based on *commercial off the shelf* (COTS) components from Sundance multiprocessor [29]. More specifically, the transmit node contains four DACs, which generate *intermediate-frequency* (IF) signals to feed the RF front-end only through the I branch (while the Q branch is not used). Given that an IF signal is provided to a direct-conversion front-end, the desired signal plus an undesired replica are recovered at its output. Nevertheless, this replica is suppressed at the receiver by shifting the RF carrier frequency and by an adequate filtering in the digital domain.

Both transmit and receive nodes make use of real-time buffers which are used to store the signals to be sent to the DACs as well as the signals acquired by the *analog-to-digital converters* (ADCs). The utilization of such buffers allows for transmission and acquisition of signals in real-time, while signal generation and processing is carried out off-line. Additionally, both baseband hardware and RF front-ends of all the nodes are synchronized in both time and frequency by means of two mechanisms:

- First, a hardware trigger is attached to the real-time buffers and to the DACs at the transmitter. When the trigger is switched on (this action can also be started by the user directly from MATLAB), all the buffers and DACs receive the trigger signal and they start to transmit simultaneously.
- Second, the same common external 40 MHz reference oscillator is utilized by both DACs and RF front-ends of all the nodes, thus guaranteeing frequency synchronization.

5.2 Measurement set-up

At the transmitter side, once the discrete-time and complex-valued sequences to be transmitted have been generated, the resulting symbols are assembled into the frame structure plotted in Figure 6. A *pseudo noise* (PN) sequence of 119 BPSK symbols is appended for frame detection and synchronization. Next, a 50 symbol duration silence period is intentionally left to estimate the Rx-noise covariance matrix. Finally, 4 000 4-QAM modulated pilot symbols are transmitted per antenna in order to estimate the MIMO channel.

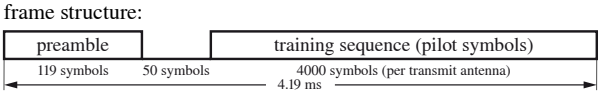

Figure 6 Frame structure used in the measurements.

The frame structure plotted in Figure 6 is suitable for estimating the covariance matrices of both receiver and transmitter noise. On the one hand, from the signals acquired during the silence period, the Rx-noise covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{\eta}_r}$ is estimated. On the other hand, both the MIMO channel $\hat{\boldsymbol{H}}$ and the total noise covariance matrix, $\hat{\boldsymbol{C}}_{\boldsymbol{\eta}}$, are estimated from the signals measured when the training sequence is transmitted. According to Equations (1) and (4), the total noise covariance matrix is $\boldsymbol{C}_{\boldsymbol{\eta}} = \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{\eta}_r}$. Hence, the estimate of the Tx-noise covariance matrix, $\hat{\boldsymbol{C}}_{\boldsymbol{\eta}_t}$, can be obtained as

$$\hat{\boldsymbol{C}}_{\boldsymbol{\eta}_t} = \hat{\boldsymbol{H}}^{-1}(\hat{\boldsymbol{C}}_{\boldsymbol{\eta}} - \hat{\boldsymbol{C}}_{\boldsymbol{\eta}_r})\hat{\boldsymbol{H}}^{-\mathrm{H}}, \tag{36}$$

where the superscript $^{-\mathrm{H}}$ denotes the inverse transpose conjugate operator.

An up-sampling by a factor of 40 (i.e., 40 samples per symbol) and a pulse-shape filtering using a squared root-raised cosine filter with 12% roll-off are performed. Consequently, given that the sampling frequency of the DACs is set to 40 MHz, the signal bandwidth is 1.12 MHz, which leads–according to our tests–to a frequency-flat channel response. Note that DACs implement an internal, interpolating filter which improves the signal quality at its output, resulting in a sampling rate of 160 Msample/s. The obtained signals are I/Q modulated to form a passband signal at a carrier frequency of f_{IF} = 3 MHz. Such signals are then properly scaled in order to guarantee the transmit-energy constraint. Moreover, given that the DACs have a resolution of 16 bits, the signals are properly quantized, giving as a result 16-bit integer values for the samples. These signals are stored off-line in buffers available at the transmit nodes of the testbed.

Afterwards, the transmitter triggers all the DACs at the same time, so that the buffers are simultaneously and cyclically read in real-time by the corresponding DACs, hence generating signals at the IF. Next, these analog signals are sent to the RF front-end to be transmitted at the desired RF center frequency at a mean transmit power of 5dBm per transmit antenna.[c] With the aim of obtaining statistically-rich channel realizations, and given that the RF front-end is frequency-agile, we measured at different RF carriers (frequency hopping) in the frequency interval ranging from 5219 MHz to 5 251 MHz and from 5 483 MHz to 5 703 MHz. Carrier spacing is 4MHz (greater than the signal bandwidth), which results in 65 different frequencies. Additionally, we repeated this measurement procedure for four different spatial positions of the receiver, giving as a result 260 different channel realizations. Channel matrices were normalized in order that the mean

Frobenius norm after averaging over the 260 channel realizations be equal to $N_t N_r = 16$.

Note that 260 channel realizations constitute a rich set of independent statistical samples from the channel (covering both 2.5 and 5GHz ISM bands) because we analyze the results in relative terms, i.e., we always compare a certain curve to a so-called reference curve and we focus our attention in the difference (improvement) rather than in the absolute values obtained. In this case, when the difference is the target of the analysis and not the absolute values, both curves (the reference and the one corresponding to the proposed system) are correlated, thus the uncertainty associated to the results is much lower than that of the individual curves. An example can be found in [30].

At the receiver side, once the transmitter has been triggered, signals are processed as follows. First, the RF front-end down-converts the signals received by the selected antennas (up to four) to the baseband, generating the corresponding analog passband signals. All the IF signals are then digitized by the ADCs by sampling at 40MHz and they are taken in realtime to the corresponding buffers. Given that the signals are being cyclically transmitted and in order to guarantee that a frame is entirely captured, all the transmit frame is acquired twice. The signals are then properly scaled according to the 14 bits resolution exhibited by ADCs. Notice that this factor is constant during the whole measurement, hence not affecting the channel properties. Next, the acquired signals are filtered using a custom-designed passband filter, which eliminates all the undesired replicas of the signals. Next, time and frequency synchronization are carried out making use of the known preamble. As soon as the acquired frames are correctly synchronized, the resulting signals are I/Q demodulated and filtered again (matched filter) so that as a result, discrete-time, and complex-valued observations with 40 samples per symbol are obtained. After filtering, the signals are decimated. Finally, the frame is properly disassembled, and final observations are then sent to the channel estimator. The raw acquired signals, i.e., the discrete-time complex-valued observations (including those obtained during the silence period shown in Figure 6), as well as the channel estimates, are finally stored for subsequent evaluation.

5.3 Results with measured indoor channels

In total, 260 different realizations were obtained from the measurement campaign.[d] The same measurement is utilized with two objectives. First, based on the frame structure shown in Figure 6, we estimate the channel coefficients. Second, making use of the silence period included in the frame, the noise variance at the receiver (Rx-noise) from each acquired frame is also estimated. Such a noise variance estimation is readily obtained by calculating the variance of the samples corresponding to the silence period after filtering and decimation at the receiver.

After averaging over all the 260 Rx-noise realizations obtained from the measurement campaign, the following (unnormalized) covariance matrix is obtained for the Rx-noise,

$$\hat{C}_{\eta_r} = 10^2 \times \begin{bmatrix} 1.2058 & -0.0104+0.0079\mathrm{i} & -0.0030+0.0141\mathrm{i} & 0.0212-0.0040\mathrm{i} \\ -0.0104-0.0079\mathrm{i} & 0.9973 & 0.0119+0.0061\mathrm{i} & 0.0043-0.0038\mathrm{i} \\ -0.0030-0.0141\mathrm{i} & 0.0119-0.0061\mathrm{i} & 0.6265 & -0.0007+0.0076\mathrm{i} \\ 0.0212+0.0040\mathrm{i} & 0.0043+0.0038\mathrm{i} & -0.0007-0.0076\mathrm{i} & 1.1025 \end{bmatrix}, \quad (37)$$

whereas the corresponding Tx-noise covariance matrix (also unnormalized) is obtained from 260 different measurements carried out connecting the transmitter and the receiver using coaxial cables plus attenuators (we repeat exactly the same procedure for measuring both the wireless channels and the noise at the transmitter). Notice that the configuration parameters of the transmitter are exactly the same as before, when measuring $\hat{C}_{\eta_r}$ by means of wireless transmissions. However, by using cables plus attenuators, such a Tx-noise covariance matrix estimation is significantly more accurate as well as more precise because the channel matrix $\boldsymbol{H}$ in Equation (36) can be assumed to be known (actually, the cabled channel is estimated with very high precision and accuracy). Such a Tx-noise covariance matrix is given by

$$\hat{C}_{\eta_t} = 10^{-3} \times \begin{bmatrix} 2.6069 & 0.0810-0.0115\mathrm{i} & 0.0058+0.0068\mathrm{i} & -0.0107+0.0001\mathrm{i} \\ 0.0081+0.0115\mathrm{i} & 2.6413 & 0.0020+0.0020\mathrm{i} & -0.0127+0.0035\mathrm{i} \\ 0.0058-0.0068\mathrm{i} & 0.0020-0.0020\mathrm{i} & 2.5552 & 0.0037+0.0018\mathrm{i} \\ -0.0107-0.0001\mathrm{i} & -0.0127-0.0035\mathrm{i} & 0.0037-0.0018\mathrm{i} & 3.4155 \end{bmatrix}. \quad (38)$$

Thus, the STxNR is estimated by performing $N_t/\mathrm{tr}(\hat{C}_{\eta_t})$, with $\hat{C}_{\eta_t}$ given by Equation (38), obtaining as a result the value of 25.52 dB, which is in accordance with the practical range between 22 dB and 32 dB.

The figure of merit chosen for the performance evaluation is the uncoded BER versus SRxNR, as defined in Equation (3). The symbols are QPSK modulated and it is also assumed that $\boldsymbol{C}_u = \boldsymbol{I}$. The channel is quasi-static and remains unchanged during the transmission of a frame of 10^5 symbols. Therefore, additive Gaussian Rx-noise and Tx-noise were artificially generated and spatially colored according to matrices of Equations (37) and (38), respectively, to recreate the measured scenario. Also note that additional noise was injected to change the operating SRxNR value of the BER curves.

Figures 7 to 10 show the results after averaging over all the 260 channel realizations obtained from this measurement campaign. Each figure plots the uncoded BER versus SRxNR curves corresponding to the following three different situations: i) system performance for the

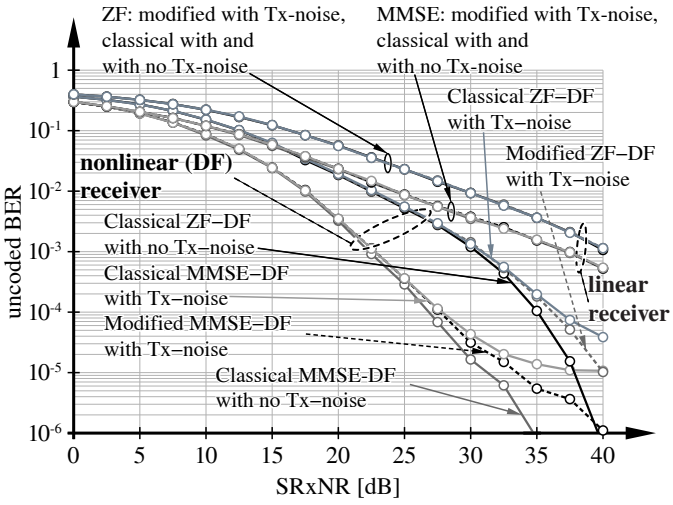

Figure 7 BER versus SRxNR with MIMO linear and nonlinear (DF) receivers over measured indoor channels. Figure 7 shows the uncoded BER with respect to the SRxNR. A MIMO system with four transmit antennas and four receive antennas ($N_t = N_r = 4$) is considered. The performance evaluation is carried out for ZF as well as MMSE, considering linear receivers and nonlinear (DF) receivers, while all schemes are evaluated with and without Tx-noise. Finally, schemes specifically designed to deal with Tx-noise (labeled as "modified") are also evaluated. In all cases, QPSK signals over measured indoor channels are used, whereas the estimated STxNR is about 25dB.

ideal case in which there is no Tx-noise (curves labeled as "Classical ... with no Tx-noise"); ii) system performance when there is Tx-noise but the system design does not take it into account (curves labeled as "Classical ... with Tx-noise"); and iii) system performance when there is Tx-noise which is considered for the system design (curves labeled as "Modified ... with Tx-noise").

Figure 7 shows the performance of the MIMO system with both MMSE and ZF linear receivers. As it can be seen from the figure, linear receivers are very little affected by the presence of Tx-noise. Figure 7 also plots the average BER versus SRxNR for MIMO systems with both MMSE DF and ZF DF receivers. Again, Tx-noise does not severely degrade system performance (except for the MMSE DF at SRxNR values greater than 30dB). Considering the Tx-noise in the DF receiver design only produces a small improvement in performance.

TX-noise, however, severely degrades the performance of MIMO systems with linear precoding (LP), as it is plotted in Figure 8. Notice how performance starts to degrade for SRxNR values above 20 dB. Tx-noise causes an error floor of about 10^{-2}, which can be reduced to 3×10^{-3} when it is considered in the MMSE precoder design.

Figure 9 plots the BER performance in terms of SRxNR for MIMO systems with both MMSE and ZF TH precoding. Note the high sensitivity of the performance of MIMO systems with THP with respect to Tx-noise, which produces that the curves go up again for high scenarios [31]. The performance of both MMSE and ZF TH precoding is considerably degraded due to the Tx-noise and an error floor arises at the significant value of 2×10^{-3} (at 40dB of SRxNR). This error floor reduces to about 2×10^{-6} (at 40 dB of SRxNR) when the Tx-noise is considered in the THP design. Therefore, the increasing effect in BER due to Tx-noise, specially for high SNR, is strongly mitigated when the design is robust against this type of impairments.

The same behaviour is observed for VP although at lower values of BER since VP performs better than THP and it is not so influenced by noise impairments at the transmitter due to the correction performed by the lattice search. Figure 10 shows how Tx-noise degrades the performance of MIMO systems with MMSE VP for SRxNR values greater than 20dB introducing an error floor at 10^{-5}. This error floor reduces to 2×10^{-6} when the Tx-noise is introduced in the MMSE VP design.

In general, in all cases–linear precoding, Tomlinson-Harashima, and vector precoding– for SRxNR values smaller than 20dB, the noise at the receiver is the dominant factor, so that STxNR values around 25dB do not influence the uncoded BER performance. Note that the STxNR value of 25.52dB determines the SRxNR value at which the performance begins to be improved by considering Tx-noise in the system design.

Figures 9 and 10 present the performance evaluation results in terms of uncoded BER versus SRxNR for TH precoding as well as VP when QPSK mappings are utilized. However, higher modulation levels lead to a greater loss in performance when the Tx-noise is not included in the designs as it can be seen in Figures 11 and 12, where the performance of Tomlinson-Harashima and vector

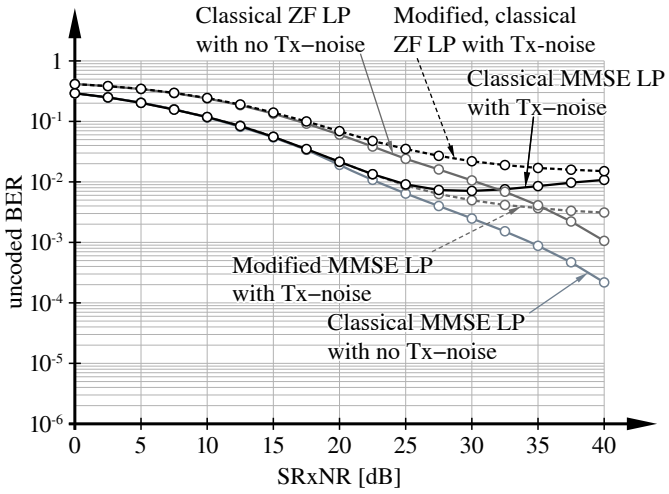

Figure 8 BER versus SRxNR with MIMO LP over measured indoor channels. Figure 8 shows the uncoded BER with respect to the SRxNR. A MIMO system with four transmit antennas and four receive antennas ($N_t = N_r = 4$) is considered. The performance evaluation is carried out considering LP for ZF as well as MMSE, while all schemes are evaluated with and without Tx-noise. Finally, schemes specifically designed to deal with Tx-noise (labeled as "modified") are also evaluated. In all cases, QPSK signals over measured indoor channels are used, whereas the estimated STxNR is about 25 dB.

precoding using 16-QAM is, respectively, shown. In both cases, the proposed modified designs will mitigate that loss produced by Tx-noise.

Finally, note also that for increasing SRxNR values, all curves corresponding to MMSE-based systems converge to those corresponding to ZF-based systems.

6 Conclusions

This study investigates the impact of transmitter noise on the performance of several MIMO systems, namely, linear and decision feedback receivers, as well as linear, Tomlinson-Harashima, and vector precoders. In all these cases, both MMSE and ZF designs were considered. We derived the expressions for all the transceiver designs taking into account the presence of transmitter noise. We also presented the results that illustrate the performance of all the aforementioned transmission schemes using MIMO channels and noise models, both obtained from an indoor measurement campaign carried out with a MIMO testbed developed at the University of A Coruña.

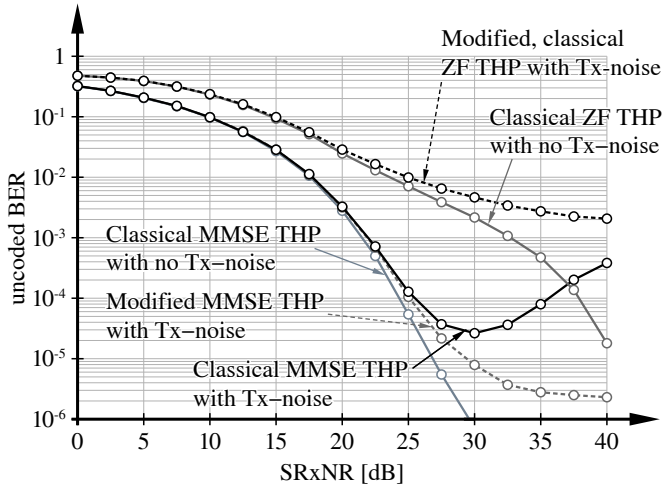

Figure 9 BER versus SRxNR with MIMO THP over measured indoor channels. Figure 9 shows the uncoded BER with respect to the SRxNR. A MIMO system with four transmit antennas and four receive antennas ($N_t = N_r = 4$) is considered. The performance evaluation is carried out considering THP for ZF as well as MMSE, while all schemes are evaluated with and without Tx-noise. Finally, schemes specifically designed to deal with Tx-noise (labeled as "modified") are also evaluated. In all cases, QPSK signals over measured indoor channels are used, whereas the estimated STxNR is about 25 dB.

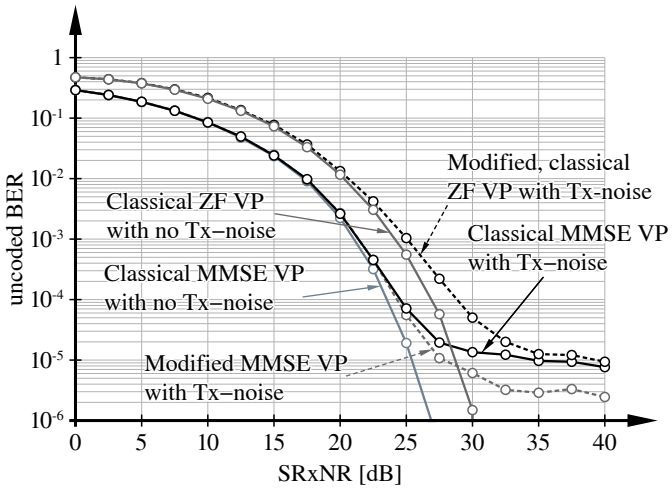

Figure 10 BER versus SRxNR with MIMO VP over measured indoor channels for $N_t = N_r = 4$. Estimated STxNR is about 25dB. Figure 10 shows the uncoded BER with respect to the SRxNR. A MIMO system with four transmit antennas and four receive antennas ($N_t = N_r = 4$) is considered. The performance evaluation is carried out considering VP for ZF as well as MMSE, while all schemes are evaluated with and without Tx-noise. Finally, schemes specifically designed to deal with Tx-noise (labeled as "modified") are also evaluated. In all cases, QPSK signals over measured indoor channels are used, whereas the estimated STxNR is about 25dB.

Basically, we conclude that transmitter noise has little impact on the performance–expressed in terms of uncoded BER versus signal-to-noise ratio at the receiver–of MIMO linear and decision feedback receivers, but marginal performance improvements can be expected when considering transmitter noise in the receiver design. On the contrary, transmitter noise considerably degrades the performance of precoded MIMO systems. Nevertheless, significant improvements on performance are obtained when transmitter noise is included in the precoder design. As expected, systems with VP produce the best results while the worst ones are obtained with LP.

Appendix 1: Details of robust optimizations

In this appendix, we will derive the expressions of all the linear and nonlinear filters shown throughout this article including the transmitter noise in the optimizations in order to obtain the MMSE robust designs appropriate to mitigate such transmitter impairments.

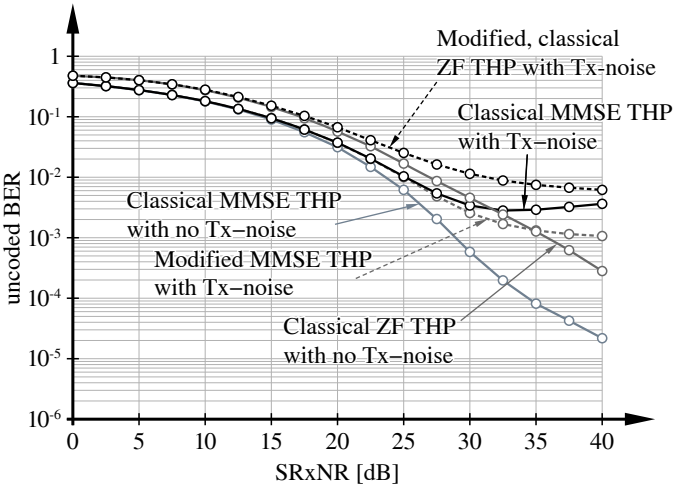

Figure 11 The same as Figure 9 but for 16-QAM instead of QPSK. Figure 11 shows the uncoded BER with respect to the SRxNR. A MIMO system with four transmit antennas and four receive antennas ($N_t = N_r = 4$) is considered. The performance evaluation is carried out considering THP for ZF as well as MMSE, while all schemes are evaluated with and without Tx-noise. Finally, schemes specifically designed to deal with Tx-noise (labeled as "modified") are also evaluated. In all cases, 16-QAM signals over measured indoor channels are used, whereas the estimated STxNR is about 25dB.

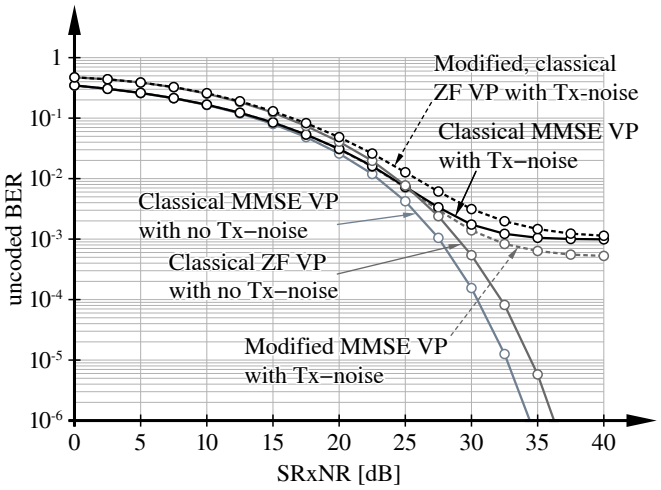

Figure 12 The same as Figure 11 but for 16-QAM instead of QPSK. Figure 12 shows the uncoded BER with respect to the SRxNR. A MIMO system with four transmit antennas and four receive antennas ($N_t = N_r = 4$) is considered. The performance evaluation is carried out considering VP for ZF as well as MMSE, while all schemes are evaluated with and without Tx-noise. Finally, schemes specifically designed to deal with Tx-noise (labeled as "modified") are also evaluated. In all cases, 16-QAM signals over measured indoor channels are used, whereas the estimated STxNR is about 25 dB.

Appendix 1.1: Derivation of MIMO linear receiver design with transmitter noise

The mean square error with Tx-noise in the MMSE linear receiver design is given by

$$\begin{aligned}\mathrm{E}\left[\left\|\varepsilon_t[n]\right\|_2^2\right] &= \mathrm{tr}\,(\boldsymbol{C_u}) - \mathrm{tr}\left(p^*\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) - \mathrm{tr}\left(p\boldsymbol{G}\boldsymbol{H}\boldsymbol{C_u}\right)\\ &\quad + \mathrm{tr}\left(|p|^2\boldsymbol{G}\boldsymbol{H}\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}_r}\boldsymbol{G}^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right),\end{aligned}$$

which can be expressed similarly to Equation (9) by doing $\mathrm{E}\left[\left\|\varepsilon[n]\right\|_2^2\right]$, i.e., the mean square error without Tx-noise, equal to $\mathrm{tr}\,(\boldsymbol{C_u}) - \mathrm{tr}\left(p^*\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) - \mathrm{tr}\left(p\boldsymbol{G}\boldsymbol{H}\boldsymbol{C_u}\right) + \mathrm{tr}\left(|p|^2\boldsymbol{G}\boldsymbol{H}\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}_r}\boldsymbol{G}^{\mathrm{H}}\right)$.

Then, we construct the Lagrangian function as follows

$$L_t(p, \boldsymbol{G}, \lambda) = E\left[\left\|\varepsilon_t[n]\right\|_2^2\right] + \lambda\left(|p|^2\mathrm{tr}\,(\boldsymbol{C_u}) - E_{\mathrm{tx}}\right),$$

with the Lagrangian multiplier $\lambda \in \mathbb{R}^{0,+}$.

We equate the derivatives with respect to p and $\boldsymbol{G}$ to zero, which leads to the following *Karush-Kuhn-Tucker* (KKT) optimality conditions [32-35]

$$\begin{aligned}\frac{\partial L_t(\bullet)}{\partial \boldsymbol{G}*} &= -p^*\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}} + |p|^2\boldsymbol{G}\boldsymbol{H}\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}_r} = 0\\ \frac{\partial L_t(\bullet)}{\partial p} &= -\mathrm{tr}\,(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C_u}) + p^*\mathrm{tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) + \lambda p^*\mathrm{tr}\,(\boldsymbol{C_u}) = 0\\ |p|^2\mathrm{tr}\,(\boldsymbol{C_u}) &\le E_{\mathrm{tx}}\\ \lambda\left(|p|^2\mathrm{tr}\,(\boldsymbol{C_u}) - E_{\mathrm{tx}}\right) &= 0 \quad \text{with } \lambda \ge 0.\end{aligned}$$

From the first equation, we obtain the following expression for the receive filter $\boldsymbol{G}$

$$\boldsymbol{G} = p^*\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\left(|p|^2\boldsymbol{H}\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}.$$

By plugging this result into the second KKT condition, it is easy to demonstrate that $\lambda > 0$, and therefore the energy transmit constraint is maintained. To ensure a unique solution, we restrict $p \in \mathbb{R}^+$. Thus, p is obtained from the energy transmit constraint and we have that $p = \sqrt{\frac{E_{\mathrm{tx}}}{\mathrm{tr}(\boldsymbol{C_u})}}$. Applying the matrix inversion lemma to the above expression for the receive filter $\boldsymbol{G}_{\mathrm{MMSE}}$, it can be demonstrated that

$$\begin{aligned}\boldsymbol{G}_{\mathrm{MMSE}} &= p_{\mathrm{MMSE}}\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{C}_{\boldsymbol{\eta}}^{-1} - \boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\left(p_{\mathrm{MMSE}}^{-2}\mathbf{I} + \boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\right)\\ &= p_{\mathrm{MMSE}}\left(\boldsymbol{C_u} - \boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\left(p_{\mathrm{MMSE}}^{-2}\mathbf{I} + \boldsymbol{C_u}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{C_u}\right)\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\\ &= p_{\mathrm{MMSE}}\left(\boldsymbol{C_u}^{-1} + p_{\mathrm{MMSE}}^2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1},\end{aligned}$$

where $\boldsymbol{C}_{\boldsymbol{\eta}} = \boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{H}$.

Then, we easily reach the solution for the MMSE linear receiver in Equation (10).

Appendix 1.2: Derivation of MIMO linear precoder design with transmitter noise

The mean square error for the MMSE linear precoder design with Tx-noise is given by

$$\begin{aligned}\mathrm{E}\left[\left\|\varepsilon_t[n]\right\|_2^2\right] &= \mathrm{tr}\,(\boldsymbol{C_u}) - \mathrm{tr}\left(g^*\boldsymbol{C_u}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) - \mathrm{tr}\left(g\boldsymbol{H}\boldsymbol{F}\boldsymbol{C_u}\right)\\ &\quad + |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) + |g|^2\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r}\right) + |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right),\end{aligned} \tag{39}$$

where it can be equated $\mathrm{E}\left[\|\varepsilon[n]\|_2^2\right] = \mathrm{tr}(C_u) - \mathrm{tr}\left(g^* C_u F^{\mathrm{H}} H^{\mathrm{H}}\right) - \mathrm{tr}\left(gHFC_u\right) + |g|^2 \mathrm{tr}\left(HFC_u F^{\mathrm{H}} H^{\mathrm{H}}\right) + |g|^2 \mathrm{tr}\left(C_{\eta_r}\right)$ according to Equation (15).

Thus, we can form the following Lagrangian function

$$L_t(\boldsymbol{F}, g, \lambda) = \mathrm{E}\left[\|\varepsilon_t[n]\|_2^2\right] + \lambda\left(\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right).$$

Setting the derivatives with respect to $\boldsymbol{F}$ and g to zero, we obtain the necessary KKT conditions

$$\begin{aligned}
\frac{\partial L_t(\bullet)}{\partial F*} &= -g^* H^{\mathrm{H}} C_u + |g|^2 H^{\mathrm{H}} H F C_u + \lambda F C_u = 0 \\
\frac{\partial L_t(\bullet)}{\partial g} &= -\mathrm{tr}\left(HFC_u\right) + g^* \mathrm{tr}\left(HFC_u F^{\mathrm{H}} H^{\mathrm{H}}\right) \\
&\quad + g^* \mathrm{tr}\left(C_{\eta_r}\right) + g^* \mathrm{tr}\left(HC_{\eta_t} H^H\right) = 0 \\
\mathrm{tr}\left(FC_u F^{\mathrm{H}}\right) &\le E_{\mathrm{tx}} \\
\lambda\left(\mathrm{tr}\left(FC_u F^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right) &= 0 \quad \text{with } \lambda \ge 0.
\end{aligned} \tag{40}$$

The gain g^* obtained from the second equation is given by

$$g^* = \frac{\mathrm{tr}(HFC_u)}{\mathrm{tr}(HFC_u F^{\mathrm{H}} H^{\mathrm{H}} + HC_{\eta_t} H^{\mathrm{H}} + C_{\eta_r})}. \tag{41}$$

Multiplying the first KKT condition by $\boldsymbol{F}^{\mathrm{H}}$ from the right and applying the trace operator, we get the following

$$g^* \mathrm{tr}\left(H^{\mathrm{H}} C_u F^{\mathrm{H}}\right) - |g|^2 \mathrm{tr}\left(HFC_u F^{\mathrm{H}} H^{\mathrm{H}}\right) = \lambda \mathrm{tr}\left(FC_u F^{\mathrm{H}}\right).$$

And now, combining this result with the expression for g^* in Equation (41) yields

$$\begin{aligned}
\lambda \mathrm{tr}(FC_u F^{\mathrm{H}}) &= \frac{\mathrm{tr}(HFC_u)}{\mathrm{tr}(HFC_u F^{\mathrm{H}} H^{\mathrm{H}} + HC_{\eta_t} H^{\mathrm{H}} + C_{\eta_r})} \mathrm{tr}(H^{\mathrm{H}} C_u F^{\mathrm{H}}) \\
&\quad - \frac{|\mathrm{tr}(HFC_u)|^2}{\mathrm{tr}^2(HFC_u F^{\mathrm{H}} H^{\mathrm{H}} + HC_{\eta_t} H^{\mathrm{H}} + C_{\eta_r})} \mathrm{tr}(HFC_u F^{\mathrm{H}} H^{\mathrm{H}}) \\
&= |g|^2 \mathrm{tr}(C_{\eta_r} + HC_{\eta_t} H^{\mathrm{H}}).
\end{aligned} \tag{42}$$

From the above result $\lambda = |g|^2 \xi_t$, and if we plug this result for λ into the first KKT condition, we get

$$F = \frac{1}{g}(H^{\mathrm{H}} H + \xi_t \mathbf{I})^{-1} H^{\mathrm{H}}. \tag{43}$$

By considering $\mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) = E_{\mathrm{tx}}$ and the above expression for $\boldsymbol{F}$, it is obtained that

$$|g|^2 = \frac{\mathrm{tr}\left(\left(H^{\mathrm{H}} H + \xi_t \mathbf{I}\right)^{-2} H^{\mathrm{H}} C_u H\right)}{E_{\mathrm{tx}}},$$

which leads to a unique solution given by Equation (16) if we restrict g to being positive real.

Appendix 1.3: Derivation of MIMO decision feedback receiver design with transmitter noise

The mean square error is calculated as

$$\begin{aligned}
\mathrm{E}\left[\|\varepsilon_{t,p}[n]\|_2^2\right] &= \mathrm{tr}\left(BPC_u P^{\mathrm{T}} B^{\mathrm{H}}\right) - \mathrm{tr}\left(BPC_u H^{\mathrm{H}} G^{\mathrm{H}}\right) - \mathrm{tr}\left(GHC_u P^{\mathrm{T}} B^{\mathrm{H}}\right) \\
&\quad + \mathrm{tr}\left(GHC_u H^{\mathrm{H}} G^{\mathrm{H}}\right) + \mathrm{tr}\left(GC_{\eta_r} G^{\mathrm{H}}\right) + \mathrm{tr}\left(GHC_{\eta_t} H^{\mathrm{H}} G^{\mathrm{H}}\right),
\end{aligned} \tag{44}$$

where $\mathrm{E}\left[\|\varepsilon_p[n]\|_2^2\right]$ of Equation (20) is given by $\mathrm{E}\left[\|\varepsilon_p[n]\|_2^2\right] = \mathrm{tr}\left(BPC_u P^{\mathrm{T}} B^{\mathrm{H}}\right) - \mathrm{tr}\left(BPC_u H^{\mathrm{H}} G^{\mathrm{H}}\right) - \mathrm{tr}\left(GHC_u P^{\mathrm{T}} B^{\mathrm{H}}\right) + \mathrm{tr}\left(GHC_u H^{\mathrm{H}} G^{\mathrm{H}}\right) + \mathrm{tr}\left(GC_{\eta_r} G^{\mathrm{H}}\right)$.

This allows us to construct the Lagrangian function

$$L_t(\boldsymbol{P}, \boldsymbol{G}, \boldsymbol{B}, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k) = \mathrm{E}\left[\|\varepsilon_t[n]\|_2^2\right] + 2\Re\left(\mathrm{tr}\left(\sum_{i=1}^{N} (\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{B} \boldsymbol{S}_i^{\mathrm{T}} - \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}}) \boldsymbol{\mu}_i\right)\right), \tag{45}$$

where the equality $\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{B} \boldsymbol{S}_i^{\mathrm{T}} = \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}}$, for $i = 1,\ldots, N$, must hold because of the unit lower triangular structure of $\boldsymbol{B}$. [e] To mathematically formulate this restriction, we included the selection matrix $\boldsymbol{S}_i$ defined as

$$\boldsymbol{S}_i = [0_{N-i+1 \times i-1}, \mathbf{I}_{N-i+1}] \in \{0, 1\}^{N-i+1 \times N}. \tag{46}$$

The Lagrangian multiplier $\boldsymbol{\mu}_i$, $i = 1,\ldots, N$, is a column vector of dimension $N - i + 1$.

By setting its derivatives with respect to $\boldsymbol{G}$ and $\boldsymbol{B}$ to zero, we obtain the following KKT conditions

$$\begin{aligned}
\frac{\partial L_t(\bullet)}{\partial \boldsymbol{G}*} &= -\boldsymbol{BPC}_{\boldsymbol{u}} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{GHC}_{\boldsymbol{u}} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{GHC}_{\boldsymbol{\eta}_t} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{GC}_{\boldsymbol{\eta}_r} = 0 \\
\frac{\partial L_t(\bullet)}{\partial \boldsymbol{B}*} &= \boldsymbol{BPC}_{\boldsymbol{u}} \boldsymbol{P}^{\mathrm{T}} - \boldsymbol{GHC}_{\boldsymbol{u}} \boldsymbol{P}^{\mathrm{T}} + \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{\mathrm{H}} \boldsymbol{S}_i = 0 \\
\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{B} \boldsymbol{S}_i^{\mathrm{T}} &= \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \quad \forall i \in \{1, \ldots, N\}.
\end{aligned} \tag{47}$$

From the second KKT condition, we obtain

$$\boldsymbol{B} = \boldsymbol{GHP}^{\mathrm{T}} - \left(\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{\mathrm{H}} \boldsymbol{S}_i\right) \boldsymbol{PC}_{\boldsymbol{u}}^{-1} \boldsymbol{P}^{\mathrm{T}}. \tag{48}$$

Plugging this expression for $\boldsymbol{B}$ into the first KKT condition, we get

$$\boldsymbol{G} = -\left(\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{\mathrm{H}} \boldsymbol{S}_i\right) \boldsymbol{PH}^{\mathrm{H}} (\boldsymbol{HC}_{\boldsymbol{\eta}_t} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r})^{-1}. \tag{49}$$

Substituting into Equation (48) we obtain

$$\boldsymbol{B} = -\left(\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{H} \boldsymbol{S}_i\right) \boldsymbol{P}\left(\boldsymbol{H}^{\mathrm{H}} \left(\boldsymbol{HC}_{\boldsymbol{\eta}_t} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}}. \tag{50}$$

Applying the restriction concerned with the unit lower triangular structure of $\boldsymbol{B}$ to the above result leads to

$$\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{B} \boldsymbol{S}_i^{\mathrm{T}} = -\boldsymbol{e}_i^{\mathrm{T}} \left(\sum_{j=1}^{N} \boldsymbol{e}_j \boldsymbol{\mu}_j^{\mathrm{H}} \boldsymbol{S}_j\right) \boldsymbol{P}\left(\boldsymbol{H}^{\mathrm{H}} \left(\boldsymbol{HC}_{\boldsymbol{\eta}_t} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} = \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}}.$$

Then, with $\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{e}_j = 0$, for $j \neq i$, and 1, otherwise, $\boldsymbol{\mu}_i^{\mathrm{H}}$ reads as

$$\boldsymbol{\mu}_i^{\mathrm{H}} = -\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\left[\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1}\boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right]^{-1}.$$

This result for $\boldsymbol{\mu}_i^{\mathrm{H}}$ gives us the following expressions for the filters $\boldsymbol{G}$ and $\boldsymbol{B}$

$$\begin{aligned}\boldsymbol{G} &= \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\left[\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1}\boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right]^{-1}\times \\ &\quad\times \boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1} \\ \boldsymbol{B} &= \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\left[\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1}\boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right]^{-1}\times \\ &\quad\times \boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1}\boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right)\boldsymbol{P}^{\mathrm{T}}.\end{aligned} \tag{51}$$

Bearing in mind the *Cholesky factorization with symmetric permutation*, the feedforward filter in Equation (51) reduces to

$$\begin{aligned}\boldsymbol{G} &= \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{D}^{-1}\boldsymbol{L}^{-1}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1} \\ &= \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{D}^{-1}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{-1}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1} \\ &= \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1} \\ &= \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1} = \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ &= \boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r}\right)^{-1},\end{aligned} \tag{52}$$

where in the derivations we have used the following properties for the selection matrix $\boldsymbol{S}_i$

$$\boldsymbol{S}_i\boldsymbol{N} = \boldsymbol{S}_i\boldsymbol{N}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i, \quad \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{M}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i = \boldsymbol{e}_i^{\mathrm{T}}, \quad \text{and} \quad \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{N}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i = \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{N},$$

with $\boldsymbol{N}$ being an upper triangular matrix and $\boldsymbol{M}$ a unit lower triangular matrix. Comparing this result with Equation (49) leads to the conclusion that $-\sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{\mu}_i^{\mathrm{H}}\boldsymbol{S}_i = \boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}$. Hence, the feedback filter reduces to

$$\boldsymbol{B} = \boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{D}^{-1}\boldsymbol{L}^{-1} = \boldsymbol{L}^{-1}. \tag{53}$$

Therefore, the filters $\boldsymbol{B}$ and $\boldsymbol{G}$ corresponding to the MMSE DF receiver solution are given by Equation (21).

Appendix 1.4: Derivation of MIMO Tomlinson-Harashima design with transmitter noise

The mean square error is given by

$$\begin{aligned}\mathrm{E}\left[\left\|\varepsilon_i[n]\right\|_2^2\right] &= \mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\upsilon}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) - g^*\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) - g\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) \\ &\quad + |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) + |g|^2\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r}\right) + |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right),\end{aligned} \tag{54}$$

where the mean square error without taking into account in the optimizations the Tx-noise is given by $\mathrm{E}\left[\left\|\varepsilon[n]\right\|_2^2\right] = \mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\upsilon}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) - g^*\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) - g\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) + |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) + |g|^2\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r}\right)$ accordingly to Equation (25).

Then, the MSE in Equation (54) enables us to construct the Lagrangian function as follows

$$\begin{aligned}L_t(\boldsymbol{P}, \boldsymbol{B}, \boldsymbol{F}, g, \lambda, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K) &= \mathrm{E}\left[\left\|\varepsilon_t[n]\right\|_2^2\right] + \lambda\left(\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right) \\ &\quad + 2\Re\left(\sum_{i=1}^{K}\mathrm{tr}\left(\boldsymbol{\mu}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i - \boldsymbol{S}_i\boldsymbol{e}_i\right)\right)\right),\end{aligned} \tag{55}$$

where $\boldsymbol{S}_i$ is a selection matrix defined as [cf. Equation (46)], $\lambda \in \mathbb{R}^{0,+}$, $\boldsymbol{\mu}_i \in \mathbb{C}^i$, $i = 1,\ldots, K$, and $2\Re\left(\sum_{i=1}^{K}\mathrm{tr}\left(\boldsymbol{\mu}_i^{T}\left(\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i - \boldsymbol{S}_i\boldsymbol{e}_i\right)\right)\right)$ comes from the restriction for the unit lower triangular structure of the feedback matrix $\boldsymbol{B}$.

Setting the derivatives of the Lagrangian function with respect to $\boldsymbol{B}$, $\boldsymbol{F}$, and g to zero we obtain

$$\begin{aligned}\frac{\partial L_t(\bullet)}{\partial \boldsymbol{F}*} &= -g^*\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\upsilon} + |g|^2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon} + \lambda\boldsymbol{F}\boldsymbol{C}_{\upsilon} = 0 \\ \frac{\partial L_t(\bullet)}{\partial \boldsymbol{B}*} &= \boldsymbol{B}\boldsymbol{C}_{\upsilon} - g\boldsymbol{P}\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon} + \sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}} = 0 \\ \frac{\partial L_t(\bullet)}{\partial g} &= -\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) + g^*\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) \\ &\quad + g^*\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{H}\right) + g^*\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r}\right) = 0 \\ \boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i &= \boldsymbol{S}_i\boldsymbol{e}_i \\ \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) &\leq E_{\mathrm{tx}} \\ \lambda\left(\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right) &= 0 \quad \text{with } \lambda \geq 0.\end{aligned} \tag{56}$$

The weight g^* resulting from the third KKT condition is expressed as

$$g^* = \frac{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right)}{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)}. \tag{57}$$

If we multiply the first KKT condition from the right by $\boldsymbol{F}^{\mathrm{H}}$ and afterwards apply the trace operator we get

$$\lambda\,\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) = g^*\mathrm{tr}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) - |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right).$$

Plugging Equation (57) into the above equation, we can easily derive that

$$\lambda\,\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) = |g|^2\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right),$$

and then, $\lambda = |g|^2\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)/\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) > 0$ if we omit the trivial solution $\boldsymbol{F} = \boldsymbol{0}$. Therefore, the transmit energy constraint is active, i.e., $\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\upsilon}\boldsymbol{F}^{\mathrm{H}}\right) = E_{\mathrm{tx}}$ and $\lambda = |g|^2\xi_t$ with $\xi_t = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_r} + \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\eta}_t}\boldsymbol{H}^{\mathrm{H}}\right)/E_{\mathrm{tx}}$, as before.

Thus, the resulting feedforward filter $\boldsymbol{F}$ obtained from the first equality in Equation (56) is given by

$$\boldsymbol{F} = \frac{1}{g}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi_t\boldsymbol{I}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B} = \frac{1}{g}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\boldsymbol{I}\right)^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}, \tag{58}$$

where we applied the matrix inversion lemma to get the last equality.

By plugging the above result into the second KKT condition, we obtain that

$$\frac{\partial L_t(\bullet)}{\partial \boldsymbol{B}*} = \boldsymbol{B}\boldsymbol{C}_v - \boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_v + \sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}}$$
$$= \xi_t\boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_v + \sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}} = 0.$$

Therefore, the feedback filter $\boldsymbol{B}$ is expressed as

$$\boldsymbol{B} = -\xi_t^{-1}\boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}}\sigma_{v,i}^{-2}, \tag{59}$$

where we included the assumption that the entries of $\boldsymbol{v}[n]$ are uncorrelated.

Multiplying this result by $\boldsymbol{S}_i$ from the left and by $\boldsymbol{e}_i$ from the right, we have

$$\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i = -\xi_t^{-1}\boldsymbol{S}_i\boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\sigma_{v,i}^{-2} = \boldsymbol{S}_i\boldsymbol{e}_i.$$

Then, the Lagrangian multipliers $\boldsymbol{\mu}_i^*$, $i = 1, \ldots, K$ are given by

$$\boldsymbol{\mu}_i^* = -\sigma_{v,i}^2\xi_t(\boldsymbol{S}_i\boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}})^{-1}\boldsymbol{S}_i\boldsymbol{e}_i. \tag{60}$$

We can now substitute $\boldsymbol{\mu}_i^*$ of Equation (60) into Equations (58) and (59) so we have the following expressions for the feedforward and feedback filters

$$\boldsymbol{F} = \frac{1}{g}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}(\boldsymbol{S}_i\boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}})^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}$$
$$\boldsymbol{B} = \boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}(\boldsymbol{S}_i\boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi_t\mathbf{I})\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}})^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}, \tag{61}$$

respectively.

Taking into account the Cholesky factorization, the precoder filters in Equation (61) can be rewritten as

$$\begin{aligned}\boldsymbol{F} &= \frac{1}{g_{\mathrm{MMSE}}^{\mathrm{THP}}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}(\boldsymbol{S}_i\boldsymbol{L}^{-1}\boldsymbol{D}^{-1}\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}})^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\\ &= \frac{1}{g_{\mathrm{MMSE}}^{\mathrm{THP}}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{D}\boldsymbol{L}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\\ &= \frac{1}{g_{\mathrm{MMSE}}^{\mathrm{THP}}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{D}\boldsymbol{L}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\\ &= \frac{1}{g_{\mathrm{MMSE}}^{\mathrm{THP}}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} = \frac{1}{g_{\mathrm{MMSE}}^{\mathrm{THP}}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\\ \boldsymbol{B} &= \boldsymbol{L}^{-1}\boldsymbol{D}^{-1}\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D} = \boldsymbol{L}^{-1},\end{aligned}$$

by considering the following properties of the selection matrix $\boldsymbol{S}_i$

$$\boldsymbol{S}_i\boldsymbol{M} = \boldsymbol{S}_i\boldsymbol{M}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i, \quad \boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{M}\boldsymbol{e}_i = \boldsymbol{e}_i \quad \text{and} \quad \boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{N}\boldsymbol{e}_i = \boldsymbol{N}\boldsymbol{e}_i, \tag{62}$$

with $\boldsymbol{M}$ being a unit lower triangular matrix and with $\boldsymbol{N}$ having an upper triangular structure. These results for $\boldsymbol{F}$ and $\boldsymbol{B}$ are shown in Equation (28).

Appendix 1.5: Derivation of MIMO vector precoder design with transmitter noise

The mean square error is given by

$$\begin{aligned}\mathrm{MSE}_{t,\mathrm{VP}} = \frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}&(\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{d}[n] - g^*\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{d}[n] - g\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{H}\boldsymbol{x}[n]\\ &+ |g|^2\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\mathrm{tr}(\boldsymbol{C}_{\eta_r}) + |g|^2\mathrm{tr}(\boldsymbol{H}\boldsymbol{C}_{\eta_t}\boldsymbol{H}^{\mathrm{H}})),\end{aligned} \tag{63}$$

where

$$\mathrm{MSE}_{\mathrm{VP}} = \frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\left(\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{d}[n] - g^*\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{d}[n] - g\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\mathrm{tr}(\boldsymbol{C}_{\eta_r})\right)$$

accordingly to Equation (33) defines the mean square error when the transmitter noise is not included in the filter optimizations.

The Lagrangian function can be expressed as

$$L_t(\boldsymbol{a}[n], \boldsymbol{x}[n], g, \lambda) = \mathrm{MSE}_{t,\mathrm{VP}} + \lambda\left(\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{x}[n] - E_{\mathrm{tx}}\right), \tag{64}$$

where $\lambda \in \mathbb{R}^{0,+}$. Now, we set its derivative with respect to $\boldsymbol{x}[n]$ and g to zero

$$\begin{aligned}\frac{\partial L_t(\bullet)}{\partial \boldsymbol{x}^*[n]} &= \frac{1}{N_{\mathrm{B}}}\left(-g^*\boldsymbol{H}^{\mathrm{H}}\boldsymbol{d}[n] + |g|^2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{x}[n]\right) + \frac{\lambda}{N_{\mathrm{B}}}\boldsymbol{x}[n] = 0\\ \frac{\partial L_t(\bullet)}{\partial g} &= \frac{1}{N_{\mathrm{B}}}\left(-\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{H}\boldsymbol{x}[n] + g^*\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{x}[n]\right)\\ &\quad + g^*\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\eta_t}\boldsymbol{H}^{\mathrm{H}}\right) + g^*\mathrm{tr}\left(\boldsymbol{C}_{\eta_r}\right) = 0\end{aligned}$$

$$\begin{aligned}&\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{x}[n] \le E_{\mathrm{tx}}\\ &\lambda\left(\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{x}[n] - E_{\mathrm{tx}}\right) = 0 \quad \text{with} \quad \lambda \ge 0.\end{aligned} \tag{65}$$

Then, the transmit symbols are directly obtained from the first KKT condition and are given by

$$\boldsymbol{x}[n] = \frac{1}{g}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \frac{\lambda}{|g|^2}\boldsymbol{I}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{d}[n]. \tag{66}$$

First of all, we have to show that $\lambda > 0$, i.e., the power constraint as active. Multiplying the second KKT condition by g, we have

$$\frac{1}{N_{\mathrm{B}}}\left(-g\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\mathrm{tr}\left(\boldsymbol{C}_{\eta_r}\right) + |g|^2\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{C}_{\eta_t}\boldsymbol{H}^{\mathrm{H}}\right)\right) = 0, \tag{67}$$

and multiplying the Hermitian of the first KKT condition by $\boldsymbol{x}[n]$ from the right, we have

$$\frac{1}{N_B}\left(-g\boldsymbol{d}^H[n]\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\boldsymbol{x}[n]^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{x}[n]\right) + \frac{\lambda}{N_B}\boldsymbol{x}^H[n]\boldsymbol{x}[n] = 0. \quad (68)$$

With Equation (67) and the transmit energy constraint, the Lagrangian multiplier λ is given by

$$\lambda = |g|^2 \frac{\mathrm{tr}\left(\boldsymbol{C}_{\eta_r} + \boldsymbol{H}\boldsymbol{C}_{\eta_t}\boldsymbol{H}^H\right)}{\frac{1}{N_B}\sum_{n=1}^{N_B}\boldsymbol{x}^H[n]\boldsymbol{x}[n]}. \quad (69)$$

Therefore, it becomes clear that $\lambda > 0$ for the non-trivial case that $\exists n : \boldsymbol{x}[n] \neq 0$. Thus, the transmit energy constraint is active, $\lambda = |g|^2 \xi_t$, and $g_{\mathrm{MMSE}}^{\mathrm{VP}}$ is directly obtained from the transmit energy constraint.

Then we reach the solution for the MMSE VP in Equation (34).

Applying the matrix inversion lemma to Equation (34) shows that $\boldsymbol{x}_{\mathrm{MMSE}}^{\mathrm{VP}}[n] = \frac{1}{g_{\mathrm{MMSE}}^{\mathrm{VP}}}\boldsymbol{H}^H\boldsymbol{\Phi}_t\boldsymbol{d}[n]$ and then, $g_{\mathrm{MMSE}}^{\mathrm{VP}} = \sqrt{\sum_{n=1}^{N_B}\left(\boldsymbol{d}^H[n]\boldsymbol{\Phi}_t^H\boldsymbol{H}\boldsymbol{H}^H\boldsymbol{\Phi}_t\boldsymbol{d}[n]\right)/\left(E_{\mathrm{tx}}N_B\right)}$. Thus, when we plug these results into the MSE expression in Equation (63) we obtain that

$$\mathrm{MSE}_{t,\mathrm{VP}} = \frac{\xi_t}{N_B}\sum_{n=1}^{N_B}\boldsymbol{d}^H[n]\boldsymbol{\Phi}_t\boldsymbol{d}[n]. \quad (70)$$

Since $\boldsymbol{\Phi}_t$ is positive definite, we can use the Cholesky factorization to obtain a lower triangular matrix $\boldsymbol{L}$ and a diagonal matrix $\boldsymbol{D}$ with the following relationship

$$\boldsymbol{\Phi}_t = \left(\boldsymbol{H}\boldsymbol{H}^H + \xi_t\mathbf{I}\right)^{-1} = \boldsymbol{L}^H\boldsymbol{D}\boldsymbol{L}.$$

Thus, the perturbation signal can be found by search in Equation (35)

Endnotes

[a]In the particular case of QPSK modulation, $\tau = 2\sqrt{2}$. [b]In our measurements, we use monopole, dual-band antennas [36]. [c]We have previously calibrated our front-ends so when we scale the passband signals prior to the DACs we ensure that the mean transmit power is fixed regardless of the RF carrier frequency utilized. We also ensure that the power amplifiers at the transmitter are operating in their linear region (note that the front-ends have an output IP3 of 25 dBm per antenna in the 5 GHz band and of 37.5 dBm in the 2.5 GHz band). [d]All measurement results (including the channel coefficients and the corresponding covariance matrices for the Tx-noise as well as for the Rx-noise are available on request). [e]The lefthand side cuts out the last N-i + 1 elements of the ith row of $\boldsymbol{B}$ and the righthand side sets the first of those elements (the ith diagonal element of $\boldsymbol{B}$) to one and the others to zero (triangularity of $\boldsymbol{B}$).

Acknowledgements
This study has been funded by Xunta de Galicia, Ministerio de Ciencia e Innovacion of Spain, and FEDER funds of the European Union under grants with numbers 10TIC003CT, 09TIC008105PR, TEC2010-19545-C04-01, and CSD2008-00010.

Competing interests
The authors declare that they have no competing interests.

Received: 20 July 2011 Accepted: 15 March 2012
Published: 15 March 2012

References

1. H Suzuki, T Tran, I Collings, G Daniels, M Hedley, Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage. IEEE J Sel Areas Commun. **26**(6), 867–876 (2008)
2. C Studer, M Wenk, A Burg, MIMO transmission with residual Tx-RF impairments, in *Proc of International ITG Workshop Smart Antennas*, Bremen, Germany, 189–196 (2010)
3. M Joham, W Utschick, JA Nossek, Linear transmit processing in MIMO communications systems. IEEE Trans Signal Process. **53**(8), 2700–2712 (2005)
4. JA Nossek, M Joham, W Utschick, Transmit processing in MIMO wireless systems, in *Proc of the 6th IEEE Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication*, Shanghai, China, **1**, I–18-I-23 (2004)
5. M Joham, Optimization of Linear and Nonlinear Transmit Signal Processing, (PhD dissertation Munich University of Technology, 2004)
6. RL Choi, RD Murch, New transmit schemes and simplified receiver for mimo wireless communication systems. IEEE Trans Wirel Commun. **2**(6), 1217–1230 (2003). doi:10.1109/TWC.2003.819024
7. HR Karimi, M Sandell, J Salz, Comparison between transmitter and receiver array processing to achieve interference nulling and diversity, in *Proc of PIMRC*, Osaka, Japan, **3**, 997–1001 (1999)
8. M Joham, K Kusume, MH Gzara, W Utschick, JA Nossek, Transmit wiener filter for the downlink of TDD DS-CDMA systems, in *Proc of ISSSTA*, Prague, Czech Republic, **1**, 9–13 (2002)
9. A Duel-Hallen, A family of multiuser decision-feedback detectors for asynchronous code-division multiple-access channels. IEEE Trans Commun. **43**(2/3/4), 421–434 (1995)
10. N Al-Dhahir, A Sayed, The Finite-Length Multi-Input Multi-Output MMSE-DFE. IEEE Trans Signal Process. **48**(10), 2921–2936 (2000). doi:10.1109/78.869048
11. RFH Fischer, *Precoding and Signal Shaping for Digital Transmission*, (John Wiley & Sons, Hoboken, New York, USA, 2002)
12. K Kusume, M Joham, W Utschick, G Bauch, Cholesky factorization with symmetric permutation applied to detecting and precoding spatially multiplexed data streams. IEEE Trans Signal Process. **55**(6), 3089–3103 (2007)
13. DA Schmidt, M Joham, W Utschick, Minimum mean square error vector precoding. Eur Trans Telecommun. **19**(3), 219–231 (2008). doi:10.1002/ett.1192
14. J González-Coma, P Castro, L Castedo, Impact of transmit impairments on multiuser MIMO non-linear transceivers, in *Proc of International ITG Workshop on Smart Antennas (WSA)*, Aachen, Germany, 1–8 (2011)
15. J González-Coma, P Castro, L Castedo, Transmit impairments influence on the performance of MIMO receivers and precoders, in *Proc of European Wireless (EW)*, Vienna, Austria, 1–8 (2011)
16. J González-Coma, P Castro, JA García-Naya, L Castedo, Performance evaluation of non-linear MIMO precoders under transmit impairments, in *Proc 19th European Signal Processing Conference*, Barcelona, Spain, (2011)
17. JA García-Naya, M González-López, L Castedo, *Radio Communications*, (INTECH 2010 chap A Distributed Multilayer Software Architecture for MIMO Testbeds)
18. FP6-IST project MASCOT, http://www.ist-mascot.org
19. A Behzad, *Wireless LAN Radios*, (John Wiley & Sons, Hoboken, New Jersey, USA, 2007)
20. C Windpassinger, RFH Fischer, T Vencel, JB Huber, Precoding in Multiantenna and Multiuser Communications. IEEE Trans Wirel Commun. **3**(4), 1305–1316 (2004). doi:10.1109/TWC.2004.830852
21. GK Kaleh, Channel equalization for block transmission systems. IEEE J Sel Areas Commun. **13**, 110–121 (1995). doi:10.1109/49.363140

22. K Kusume, M Joham, W Utschick, MMSE block decision-feedback equalizer for spatial multiplexing with reduced complexity, in *Proc IEEE Global Telecommunications Conference*, Dallas, Texas, USA, **4**, 2540–2544 (2004)
23. CP Schnorr, M Euchnerr, Lattice basis reduction: improved practical algorithms and solving subset sum problems. Math Program. **66**, 188–191 (1994)
24. JA García-Naya, O Fresnedo, FJ Vázquez-Araújo, M González-López, L Castedo, J Garcia-Frias, Experimental evaluation of analog joint source-channel coding in indoor environments, in *Proc IEEE International Conference on Communications (ICC)*, Kyoto, Japan, 1–5 (2011)
25. O González, D Ramírez, I Santamaría, JA García-Naya, L Castedo, Experimental validation of interference alignment techniques using a multiuser MIMO testbed, in *Proc International ITG Workshop on Smart Antennas (WSA)*, Aachen, Germany, 1–8 (2011)
26. Lyrtech Inc, (2011), http://www.lyrtech.com
27. Maxim Integrated Products, Inc, (2011), http://www.maxim-ic.com/
28. Ettus Research, LLC, (2011), http://www.ettus.com
29. Sundance Multiprocessor, Ltd, (2011), http://www.sundance.com
30. S Caban, JA García Naya, M Rupp, Measuring the physical layer performance of wireless communication systems: part 33 in a series of tutorials on instrumentation and measurement. IEEE Instru Meas Mag. **14**(5), 8–17 (2011)
31. M Joham, P Castro, L Castedo, W Utschick, Robust Precoding with Bayesian Error Modeling for Limited Feedback MU-MISO Systems. IEEE Trans Signal Process. **58**(9), 4954–4960 (2010)
32. W Karush, Minima of Functions of Several Variables with Inequalities as Side Conditions. MS Thesis (The University of Chicago 1939)
33. R Fletcher, *Practical Methods of Optimization*, 2nd edn. (John Wiley & Sons, UK, 2000)
34. DG Luenberger, *Linear and Nonlinear Programming*, 2nd edn. (Kluwer Academic Publishers, USA, 2004)
35. HW Kuhn, AW Tucker, Nonlinear Programming, in *Proc 2nd Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman, (University of California Press. Berkeley, CA, USA 1951) pp. 481–492
36. L-Com Antenna No. HG2458RD-SM (2010), http://www.l-com.com/item.aspx?id=22199

doi:10.1186/1687-1499-2012-109
Cite this article as: Castro *et al.*: Performance of MIMO systems in measured indoor channels with transmitter noise. *EURASIP Journal on Wireless Communications and Networking* 2012 **2012**:109.